# Structure-Function Relationships in Matrix Metalloproteinase-1

Louise Butt

The thesis is submitted in partial fulfillment of the requirements for the award of the degree of Doctor of Philosophy of the University of Portsmouth

Institute of Biomedical and Biomolecular Sciences
School of Biological Sciences
University of Portsmouth

August 2013

# ABSTRACT

Collagenolysis, the catabolism of triple helical collagen, is essential for the physiological remodeling of connective tissues during growth and development. Aberrant collagen degradation is a feature of both inflammatory diseases and cancer cell invasion. Matrix metalloproteinase-1 (MMP-1) is a known collagenase and previous studies have implicated its hemopexin (HPX) domain in binding and possibly destabilising collagen in preparation for hydrolysis by the catalytic (CAT) domain. More recently, conformational freedom and domain separation of the CAT and HPX domains has been proposed to play a role in collagen degradation. This study aims to explore HPX mediated collagen recognition and the postulated flexible state of MMP-1, in order to enhance our understanding of the collagenolytic mechanism.

Here, biophysical methods have been used to study the complex formed between the MMP-1 HPX domain and a synthetic triple helical peptide (THP) that encompasses the MMP-1 cleavage site of the collagen α1(I) chain. A programme of site-directed mutagenesis was used to produce an extensive library of recombinant proteins. Surface plasmon resonance (SPR) has been used to characterise a previously unknown collagen binding site (exosite) located in blade 1 of the HPX domain and small angle x-ray scattering has been used to probe the conformational freedom and transient domain separation in mutant forms of both zymogen and mature MMP-1 enzymes.

Significant reductions in THP affinity were observed on mutation of either Phe301, Val319 and Asp338, residues forming part of a “ball and socket” joint in the CAT-HPX interface. Disruption of the CAT-HPX interface by mutagenesis, of any of these three residues, severely impacts collagen recognition. Conversely, the reduced collagen binding activity of a "stapled" mutant (in which the CAT and HPX domains are constrained with a disulphide bridge) indicates that the ability of the domains to transiently dislocate is also important. Thus, a balanced equilibrium between these compact and dislocated states is an essential feature of MMP-1 collagenolytic activity.

# CONTENTS

# DECLARATION

Whilst registered as a candidate for the above degree I have not been registered for any other research award. The results and conclusions embodied in this thesis are the work of the named candidate and have not been submitted for any other academic award.

# LIST OF TABLES

## CHAPTER 1



## CHAPTER 2



## CHAPTER 3



## CHAPTER 4



## CHAPTER 6

# LIST OF FIGURES

## CHAPTER 1



## CHAPTER 2

## CHAPTER 3



## CHAPTER 4

## CHAPTER 5



## CHAPTER 6

## CHAPTER 7

# ABBREVIATIONS

| | |
|---|---|
| Å | Ångström (s) |
| A | absorbance |
| ADAM | a disintegrin and metalloproteinase |
| ADAMTS | a disintegrin and metalloproteinase with thromospondin motifs |
| AFM | atomic force microscopy |
| AP-1 | activator protein-1 |
| APMA | 4-aminophenyl mercuric acetate |
| APS | ammonium persulphate |
| a.u. | arbitrary unit |
| AU | absorbance unit (s) |
| α2M | α2 macroglobulin |
| βME | β-mercaptoethanol |
| BMP-1 | bone morphogenetic protein-1 |
| BSA | bovine serum albumin |
| *c* | concentration |
| CAT | catalytic domain |
| cDNA | complementary DNA |
| cfu | colony forming unit |
| CM | carboxy methyl |
| $D_2O$ | deuterium oxide |
| Da | dalton (s) |
| $ddH_2O$ | double-distilled water |
| Dmax | maximum dimension |
| DMSO | dimethyl-sulfoxide |
| DNA | deoxyribonucleic acid |
| dNTP | deoxynucleotriphosphate |
| DTNB | 5,5'-Dithiobis(2-nitrobenzoic acid) |
| DTT | dithiothreitol |
| ECM | extracellular matrix |
| *E. coli* | *Escherichia coli* |
| EDC | 1-ethyl-3 carbodiimide |

| | |
|---|---|
| EDTA | ethylenediaminetetraacetic acid |
| ELISA | enzyme-linked immunosorbent assay |
| ERK | extracellular signal-related kinase |
| F2 | fibronectin type II domain |
| FACIT | fibril-associated containing interrupted triple helices |
| FID | free induction decay |
| g | gram (s) |
| GAG | glycosaminoglycans |
| GPCR | G-protein coupled receptor |
| GPI | Glycosylphosphatidylinositol |
| GT | galacto/glucosyltransferase |
| h | hour (s) |
| HD/MS | hydrogen deuterium exchange mass spectrometry |
| HEPES | 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid |
| HPX | hemopexin domain |
| HSQC | heteronuclear single quantum coherence |
| I | intensity |
| IB | inclusion bodies |
| IL-1 | interleukin-1 |
| IMAC | immobilised metal affinity chromatography |
| $I_0$ | incident light |
| IPTG | isopropyl-β-D-thiogalactopyranoside |
| JAK | Janus kinase |
| JNK | c-Jun N-terminal activating kinase |
| $k_{cat}$ | catalytic rate |
| $k_a$ | association rate |
| $k_d$ | dissociation rate |
| $K_d$ | dissociation constant |
| kDa | kilo-Dalton (s) |
| keV | kiloelectron volt (s) |
| $k_i$ | incidence beam |
| $K_i$ | inhibitor constant |
| $K_m$ | Michaelis constant |

| | |
|---|---|
| $k_s$ | scattered beam |
| L | litre (s) |
| LB broth | Luria-Bertani broth |
| LC/MS | liquid chromatography electrospray ionization mass spectometry |
| LH | lysyl hydroxylase |
| M | molar |
| MAPK | mitogen-activated protein kinase |
| mAU | milli absorbance unit (s) |
| MD | molecular dynamics |
| MCP | monocyte chemoattractant protein |
| MES | 2-(N-morpholino)ethanesulfonic acid |
| mg | milligram (s) |
| min | minute (s) |
| mL | millilitres (s) |
| mm | millimetre |
| mM | millimolar |
| MM | molecular mass |
| MMP | matrix metalloproteinase |
| MPO | myeloperoxidase |
| mRNA | messenger RNA |
| MT-MMP | membrane-type MMP |
| MW | molecular weight |
| MWCO | molecular weight cut-off |
| NC | non-collagenous |
| n.d. | not determined |
| NF-**κ**B | nuclear factor-**κ**B |
| ng | nanogram (s) |
| NHS | N-hydroxysuccinimide |
| nm | nanometer (s) |
| nM | nanomolar |
| NMR | nuclear magentic resonance |
| NOE | nuclear Overhauser effect |
| NSD | normalised spatial discrepancy |

| | |
|---|---|
| nt | nucleotide (s) |
| NTA | nitriloacetic acid |
| OD | optical density |
| ORF | open reading frame |
| P4H | prolyl-4-hydroxylase |
| PAGE | polyacrylamide gel electrophoresis |
| PCR | polymerase chain reaction |
| PDB | protein data bank |
| PDI | protein disulphide isomerase |
| PEA-3 | polyoma enhancer A3 |
| PEG | polyethylene glycol |
| pI | isoelectric point |
| PKK | plasma kallikrein |
| Plg | plasmin |
| PMSF | phenylmethylsulfonyl fluoride |
| ppm | parts per million |
| *P(q)* | form factor |
| *P(r)* | distance distribution function |
| PRO | inhibitory pro domain |
| psi | pounds per square inch |
| *q* | reciprocal space momentum transfer vector |
| rad | radian (s) |
| R | response |
| $R_{max}$ | maximum response |
| RECK | reversion-inducing-cysteine-rich protein with kazal motifs |
| $R_g$ | radius of gyration |
| RMSD | root mean square deviation |
| RNA | ribonucleic acid |
| rpm | revolutions per minute |
| RTK | receptor tyrosine kinase |
| RU | response unit |
| SAXS | small angle x-ray scattering |
| s | second (s) |
| S.D. | standard deviation |

| | |
|---|---|
| SDM | site-directed mutagenesis |
| SDS | sodium dodecyl sulphate |
| SPR | surface plasmon resonance |
| *S(q)* | structure factor |
| STAT | signal transducer and activator of transcription |
| T7RNAP | T7 RNA polymerase |
| T | tesla (s) |
| TEMED | N, N, N', N' - tetramethylethylenediamine |
| TGF-β | transforming growth factor-β |
| THP | triple helical peptide |
| TIMP | tissue inhibitor of metalloproteinase |
| TM | transmembrane |
| TNB | 2-nitro-5-thiobenzoate |
| Tris | Tris (hydroxymethyl)-aminomethane |
| tRNA | transfer RNA |
| TROSY-HSQC | transverse relaxation-optimized spectroscopy-HSQC |
| tPA | tissue-type Plasminogen Activator |
| uPA | urokinase-type plasminogen activator |
| uPAR | urokinase-type plasminogen activator receptor |
| U | unit (s) |
| µg | microgram (s) |
| µl | microlitre (s) |
| µM | micromolar |
| UV | ultraviolet |
| V | volt (v) |
| v | volume |
| Vis | visible light |
| WT | wild-type |
| X-gal | 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside |
| θ | scattering angle |
| ε | extinction coefficient |
| λ | wavelength |
| $\rho$(r) | electron density distribution |

# ACKNOWLEDGEMENTS

As the end of my PhD journey reaches its climax, I finally have time to reflect on the past 4 years. As with all things, there have been times of adversity and frustration. Equally, there have been moments of experimental euphoria. My success stems only from the help, support and advice of those around me.

Firstly, I would like to thank my supervisor, Dr Andy Pickford. You have been patient and understanding throughout this process. I count myself fortunate and privileged to have spent the last few years under your tutelage. I would also like to thank Dr Anastasia Callaghan for her unwavering support, especially during the last year. I look forward to new and productive scientific adventures in the Callaghan Group.

My lab experience has been enriched by knowing a truly remarkable group of people. To all of my 'lab buddies' past and present, I will always be grateful. Thanks to Dr Steve Prior and Dr Laurence Arnold for their advice and guidance in the early days. Special thanks must go to Dr Jim Youell and Dr Luke Evans. You two, kept me smiling and caffeine-fuelled! I must not forget Dr Helen Vincent, for all your SAXS advice (PRIMUS is evil), Dr Katie Mumford, who donated time to proof-reading some of the chapters within this thesis, and Dr Chris Read for his generous help with NMR-related issues. A special shout-out goes to the Callaghan group. You guys kept me well fed (cookies are this PhD student's guilty pleasure) and wouldn't let my motivation wane. I would also like to thank Darren, Martin, Rich, Holly and John. It has been a joy to work with you all. If I have forgotten anyone, I apologise unreservedly.

I would like to thank all of my family and friends. Your love and support has been of great comfort to me. I must thank Rob, who never lost faith in my abilities, and Nick, for practical help and wisdom. In particular, I am indebted to my grandmother, who inspired this journey. This thesis has been a defining feature of my life for a long time. As my studies draw to a close, I hope to make up for all those lost moments.

Finally, I would like to dedicate this thesis to my parents and my son Jack. Dad, you instilled in me a great sense of purpose and direction. Mum, you taught me extraordinary resilience, tenacity and courage. Jack, you kept me balanced and grounded. Saying thank you will never be enough.

# DISSEMINATION

## Publications

The Interface between Catalytic and Hemopexin Domains in Matrix Metalloproteinase-1 Conceals a Collagen Binding Exosite

Laurence H. Arnold[§], Louise E. Butt[§], Stephen H. Prior[§], Christopher M. Read, Gregg B. Fields, and Andrew R. Pickford

[§]Joint first authorship

Abstract:

Matrix metalloproteinase-1 (MMP-1) is an instigator of collagenolysis, the catabolism of triple helical collagen. Previous studies have implicated its hemopexin (HPX) domain in binding and possibly destabilizing the collagen substrate in preparation for hydrolysis of the polypeptide backbone by the catalytic (CAT) domain. Here, we use biophysical methods to study the complex formed between the MMP-1 HPX domain and a synthetic triple helical peptide (THP) that encompasses the MMP-1 cleavage site of the collagen α1(I) chain. The two components interact with 1:1 stoichiometry and micromolar affinity via a binding site within blades 1 and 2 of the four-bladed HPX domain propeller. Subsequent site-directed mutagenesis and assay implicates blade 1 residues Phe301, Val319, and Asp338 in collagen binding. Intriguingly, Phe301 is partially masked by the CAT domain in the crystal structure of full-length MMP-1 implying that transient separation of the domains is important in collagen recognition. However, mutation of this residue in the intact enzyme disrupts the CAT-HPX interface resulting in a drastic decrease in binding activity. Thus, a balanced equilibrium between these compact and dislocated states may be an essential feature of MMP-1 collagenase activity.

## Oral presentations

1. Collagen Recognition by Matrix Metalloproteinase-1

Presented by L. Butt at the South West Structural Biology Consortium Meeting (SWSBC), University of Southampton (2012) and at the Institute of Biomolecular and Biomedical Sciences (IBBS) conference, University of Portsmouth (2012).

## Poster presentations

1. The Role of Interdomain Interactions in Matrix Metalloproteinase Latency

Presented by L. Butt at the South West Structural Biology Consortium Meeting (SWSBC), University of Reading (2011). Awarded first prize.

2. The Interface between Catalytic and Hemopexin Domains in MMP-1 Conceals a Collagen Binding Exosite

Presented by Dr A. Pickford at the Gordon Research Conference on Matrix Metalloproteinases, Bryant University, Rhode Island (2011).

# CHAPTER 1: PROTEOLYSIS OF THE EXTRACELLULAR MATRIX

## 1.1 The Extracellular Matrix

Extracellular matrices (ECMs) are complex assemblies of proteoglycans, glycosaminoglycans (GAGs) and protein components, such as laminin, elastin and collagen. These macromolecules contribute to the unique composition and topology of all tissues within multicellular vertebrate organisms, promoting both tissue organisation and structural integrity.

GAGs are a significant class of unbranched polysaccharide chains characterised by typical repeating disaccharide units, typically a hexauronic acid linked to an *N*-acetylated hexosamine, which is usually sulfated. They are ubiquitously distributed at the cell-ECM interface of eukaryotic cells and are usually found covalently linked to proteins in the form of proteoglycans. Proteoglycans are highly diverse in both form and function, from the large supramolecular structure of aggrecan bound to hyaluronic acid that keeps the matrix hydrated, to membrane-bound syndecan with both matrix adhesion and cell-signalling properties (Reviewed in Hardingham and Fosang 1992, Sasisekharan, Raman et al. 2006). The ECM is further enriched by the presence of a number of fibrillar proteins which possess structural and adhesive functions via cell surface receptors (e.g. integrins) that control tissue cohesion.

It is estimated that approximately 300 genes encode the core ECM constituents with additional ECM-associated proteins contributing to the complete 'matrisome' (Naba, Clauser et al. 2012). Not only do these diverse matrisome constituents provide structural support but act collectively to influence cellular behaviour.

Homeostasis in resting tissue is maintained through continual synthesis and deposition of these ECM constituents opposed by regulated proteolysis, while essential developmental events and wound healing require radical remodelling of connective tissues. Degradation of the ECM permits exposure and release of latent sequested bioactive molecules, such as growth factors and cytokines with causative downstream effects on cell adhesion and migration, cell surface receptor localization and instigation of signalling cascades (Reviewed in Lukashev and Werb 1998). Matrix degradation observed in both homeostasis and remodelling, is orchestrated primarily by members of

the matrix metalloproteinase (MMP) family, of which 23 homologous members have been identified in humans (Visse and Nagase 2003).

## 1.2 Collagen

Collagens are multi-gene extracellular proteins abundant throughout vertebrate tissues, providing essential structural resilience within the ECM. Currently, 28 types of collagen with 43 unique polypeptide α chains have been identified in vertebrates, alongside numerous proteins with triple helical collagenous domains (Khoshnoodi, Cartailler et al. 2006). They have been classified accordingly as fibril-forming, fibril-associated containing interrupted triple helices (FACIT), beaded filament, anchoring fibril, network-forming and transmembrane collagen (Reviewed in Ricard-Blum 2011).

All family members possess the characteristic triple helical domain but differ in size, function and tissue distribution, therefore, it is beyond the scope of this thesis to review each group. Instead further discussions will be limited to the quarter-staggered fibril-forming collagens exemplified by type I collagen.

### 1.2.1 Collagen Biosynthesis

Fibril-forming collagens are ubiquitously distributed throughout connective tissue, skin and bone and are synthesised by stromal cells, principally fibroblasts, as soluble procollagens with N and C terminal non-helical extensions (Figure 1.1).

Initially individual pro-α peptide chains are deposited into the rough endoplasmic reticulum prior to removal of the signal peptide. Here, the pro-α chains undergo significant post-translational modifications which include hydroxylation of selected prolines and lysines by the action of prolyl-4-hydroxylase (P4H) and lysyl hydroxylases (LH), glycosylation of lysine and hydroxylysine residues by galactosyl- and glucosyltransferases (GT), and sulfation of tyrosine residues (Harwood, Grant et al. 1974, Harwood, Grant et al. 1975, Kellokumpu, Sormunen et al. 1994, Passoja, Rautavuoma et al. 1998, Valtavaara, Szpirer et al. 1998). The extent of prolyl hydroxylation appears to be species dependent (Reviewed in Myllyharju 2003).

The non-collagenous C-terminal domains (NC-C) are essential for chain selection and initiation of trimeric procollagen formation (Lees, Tasab et al. 1997, Bourhis, Mariano et al. 2012). These globular domains are stabilised by the action of protein disulphide isomerase (PDI) which introduces intrachain disulphide bonds.

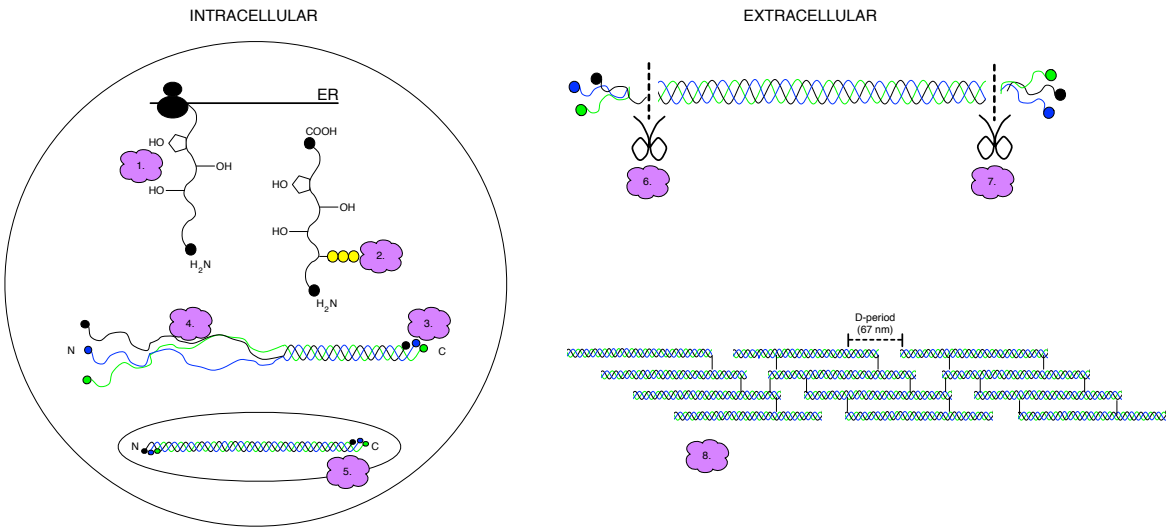



**Figure 1.1: Overview of Collagen Biosynthesis.** (**1**) Synthesis of pro α chain and hydroxylation of proline and lysine residues. (**2**) Glycosylation of lysine and hydroxylysine residues by galactosyl- and glucosyltransferases. (**3**) Nucleation at the C-terminal domains mediated by protein disulphide isomerase, and subsequent self-assembly of 3 pro α chains. (**4**) Folding promoted by peptidyl-prolyl *cis-trans* isomerase and HSP47. (**5**) Secretion of collagen monomers into the extracellular space. (**6**) Cleavage of N-terminal domains by N-proteinases, typically ADAMTs. (**7**) C-terminal processing by BMP-1. (**8**) Fibril self-assembly stabilised by lysine-derived crosslinks.

Recognition and association of the NC-C domains from three pro-α chains initiates trimerization, thus triggering propagation of the triple helix from the C-N terminus (Doege and Fessler 1986). Efficient folding is promoted by the action of peptidyl prolyl *cis-trans* isomerase which converts *cis*-peptide bonds to the *trans* form (Bachinger, Bruckner et al. 1978, Bachinger 1987, Davis, Boswell et al. 1989) and molecular chaperones like HSP47 and BiP (Chessler and Byers 1993, Sauk, Smith et al. 1994, Satoh, Hirayoshi et al. 1996, Nagai, Hosokawa et al. 2000).

The newly formed procollagen traverses the secretory pathway through the golgi body by cisternal maturation (Bonfanti, Mironov et al. 1998) and is extruded into the ECM through specialised plasma membrane structures termed 'fibripositors' (Canty, Lu et al. 2004). Procollagen is further processed in the ECM with cleavage of both NC-C and NC-N terminal domains by bone morphogenetic protein-1 (BMP-1) and closely related members of the mammalian tolloid-like metalloproteinase family, and ADAMTS (a disintegrin and metalloproteinase with thrombospondin motifs) family members displaying aminoprocollagen peptidase activity (Leung, Fessler et al. 1979, Kadler, Hojima et al. 1987, Kessler, Takahara et al. 1996, Li, Sieron et al. 1996, Fernandes, Hirohata et al. 2001, Colige, Vandenberghe et al. 2002, Wang, Lee et al. 2003).

Removal of the N and C propeptides generates tropocollagen with diminished solubility that undergoes spontaneous fibril self-assembly (Kadler, Hojima et al. 1987).

### 1.2.2 Collagen Structure

The triple helical conformation of the type I tropocollagen structural unit results from an unusual abundance of glycine (G), and imino acids proline (P) and hydroxyproline (O). Together they form the characteristic G-X-Y triplet (Figure 1.2).

The role of glycine cannot be underestimated. Tight packing of the polypeptide chains into a super-coiled helical assembly, with $10_3$ screw symmetry (3.3 residues per turn), necessitates the presence of the smallest amino acid, glycine, at every third position (Ramachandran and Kartha 1955, Rich and Crick 1955). Substitution mutations that replace invariant glycine with larger residues are known to significantly distort normal propagation, stability and function of the triple helix as seen in many types of osteogenesis imperfecta (Beck, Chan et al. 2000), a connective tissue disorder characterised by bone fragility.

Triple helical stability is greatly improved with increasing imino acid content as observed in studies of triple helical model peptides (Bella, Eaton et al. 1994, Persikov, Ramshaw et al. 2005). The restricted rotation placed on the iminopeptide bond ensures optimal chain geometry for formation of a right-handed helix with a ladder of interchain hydrogen bonds between the amide hydrogen atom of glycine and the carbonyl oxygen atom of the residue in position X in the adjacent chains (Ramachandran and Kartha 1955, Rich and Crick 1955). Stability is further enhanced by hydroxyproline in position Y, which mediates hydrogen bonding with water and induces formation of a structured hydation shell around the collagen helix (Bella, Brodsky et al. 1995).

GPO is the most commonly occurring sequence accounting for 10.5% of all triplets found in collagen. However, more than 60% of triplets distributed throughout collagen have variant amino acid residues at the Y position (Ramshaw, Shah et al. 1998). Clusters of such triplets constitute regions that are considered “imino poor”. These localised conformationally labile regions may be necessary for physiological functions. Crystallographic data shows collagen recognition by both Discoidin Domain receptor 2 and Integrin $\alpha 2\beta 1$ at imino poor sequences (Emsley, Knight et al. 2000, Carafoli, Bihan et al. 2009). Furthermore, the collagenase cleavage sites of collagen types I-III are located within an imino acid poor region (Xiao, Addabbo et al. 2010).

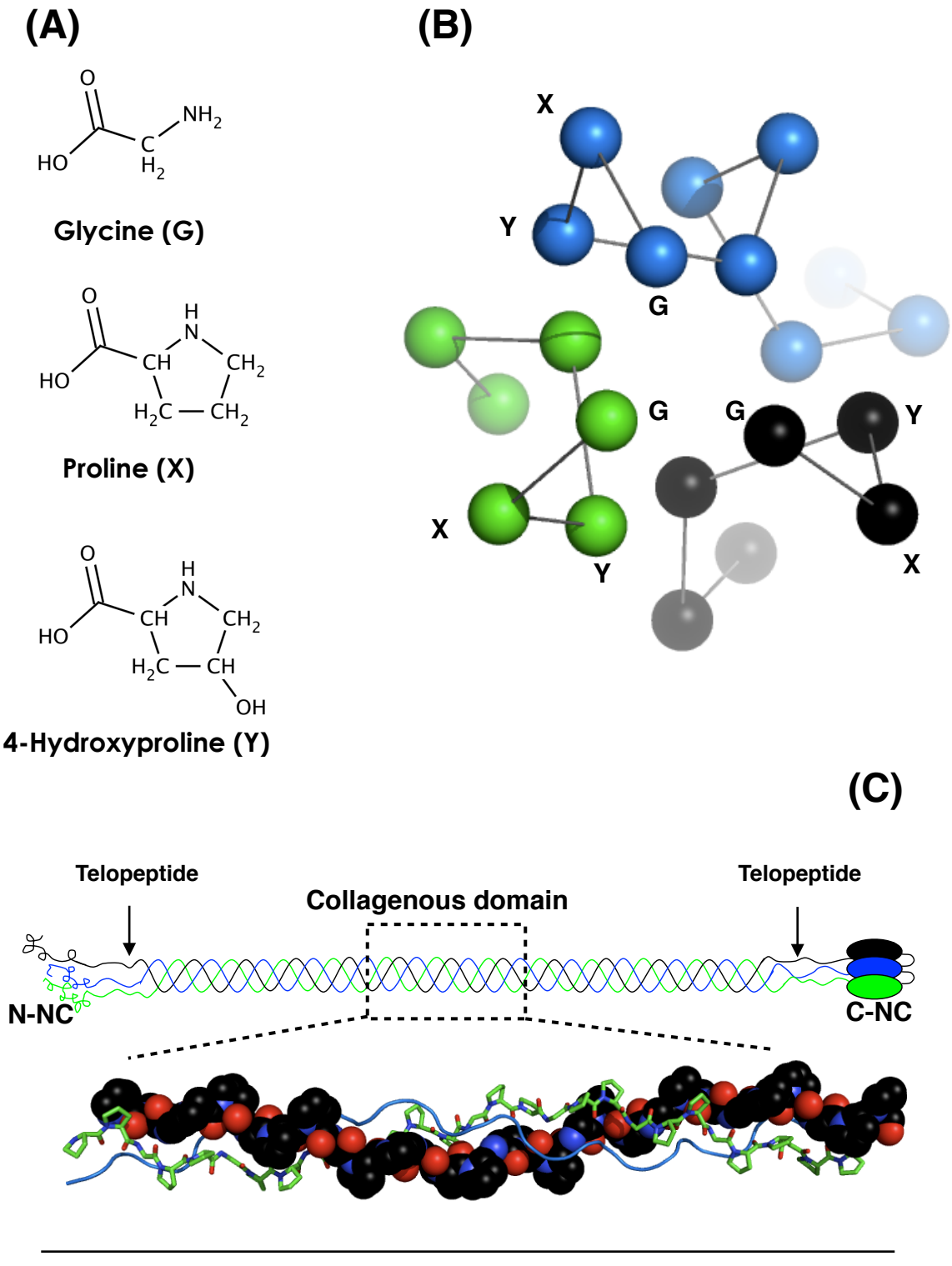



**Figure 1.2: Collagen structure.** **(A)** Structures of glycine, proline and 4-hydroxyproline. These amino residues form the characteristic G-X-Y tripeptide repeat. Structures were rendered using MarvinSketch (version 5.10.3.), ChemAxon. **(B)** Cross-sectional schematic showing the G-X-Y sequence for each α-helix. The presence of glycine at every third position ensures occupation throughout the centre of the triple helix. **(C)** Longitudinal view of a triple helical collagen monomer with sequence $[(\text{Pro-Pro-Gly})_{10}]_3$; space-filling model of chain in *black*, cartoon model of chain in *blue* and stick model of chain in *green* (1K6F.pdb). Rendered using the Pymol Molecular Graphics System (version 1.5.0.4.), Schrödinger, LLC. General domain structure is shown in schematic form, (N-NC and C-NC: N-terminal and C-terminal domains respectively).

Periodic overlapping by collagen monomers produces a 67 nm axial quarter-staggered pentameric microfibril, which is visibly striated when stained (Petruska and Hodge 1964, Baselt, Revel et al. 1993, Gutsmann, Fantner et al. 2004). These polymerized homo- and heterotrimeric collagen microfibrils are stabilised by the formation of covalent intermolecular cross links, which ultimately result from the hydroxylation pattern of lysine residues. Specific lysyl and hydroxylysyl residues in the telopeptides undergo oxidative deamination of ε-amino groups mediated by copper-dependent lysyl oxidase forming allysine and hydroxyallysine (Pinnell and Martin 1968, Siegel 1974). Immature ketoimine cross links are formed by condensation of these reactive aldehydes with lysine and hydroxylysine in helical regions of adjacent tropocollagen molecules primarily at the overlap zone of the quarter-staggered fibrils (Eyre, Weis et al. 2008). Further lateral packing of interdigitated microfibrils generates bundles of collagen fibres (Perumal, Antipova et al. 2008).

Fibrillar collagens are noted for their unique strength and resilience. The combination of structural features including a highly invariant glycine content, abundance of imino acids and extensive post-translational modifications are essential for providing protection from proteolysis, particularly by serine and sulfhydryl proteases at physiological pH (Drake, Davison et al. 1966, Bruckner and Prockop 1981, Birkedal-Hansen, Taylor et al. 1985).

## 1.3 Collagenolysis

Controlled collagen degradation and necessary remodelling within the ECM is required to ensure normal connective tissue homeostasis. Collagenolytic activity has been implicated in the physiological processes of epithelial cell migration and reepithelialisation during wound repair, formation of the embryonic circulatory system by angiogenesis, postpartum uterine involution and weakening of the preovulatory follicle wall during ovulation (Butler, Zhu et al. 1991, Alexander, Hansell et al. 1996, Seandel, Noack-Kunnmann et al. 2001).

Collagenolytic processing of interstitial collagens requires the action of collagenases; either cathepsin K, a lysosomal cysteine protease that operates in acidic environments (Garnero, Borel et al. 1998, Kafienah, Bromme et al. 1998); or certain members of the MMP family, in which an ATP-independent hydrolytic reaction occurs at neutral pH. In humans these members have been identified as MMP-1, -2, -8, -13, -14 and -16 (Gross and Lapiere 1962, Hibbs, Hasty et al. 1985, Hasty, Jeffrey et al. 1987,

Aimes and Quigley 1995, Knauper, Lopez-Otin et al. 1996, Pei and Weiss 1996, Ohuchi, Imai et al. 1997, Shimada, Nakamura et al. 1999). However, they display distinctly variable cleavage specificities. Collagen specificities have been well-described for MMP-1 (type III collagen), -8 (type I collagen) and -13 (type II collagen) (Welgus, Jeffrey et al. 1981, Hasty, Jeffrey et al. 1987, Knauper, Lopez-Otin et al. 1996, Mitchell, Magna et al. 1996).

Hydrolysis of type I-III collagen by MMPs (Figure 1.3) occurs between the Gly775-Ile776 or Gly775-Leu776 scissile bond generating characteristic ¼ and ¾ length fragments which retain triple helical structure but are unstable and undergo further enzymatic processing by gelatinases (Gross and Nagai 1965, Evanson, Jeffrey et al. 1968, Seltzer, Adams et al. 1981, Allan, Docherty et al. 1995).

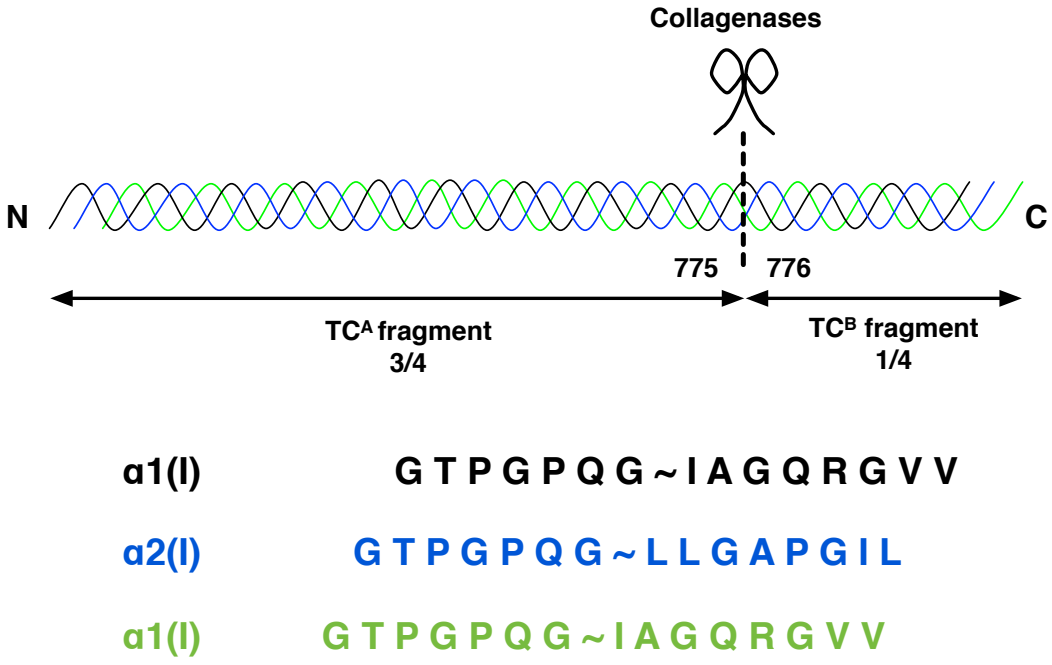



**Figure 1.3: Collagenolysis by collagenases produces characteristic ¼ and ¾ length fragments.** The heterotrimeric type I collagen cleavage site is located within the sequence 769-783 of the α1 and α2 chains (UniProt accession number P02452 and P08123 respectively) and is indicated by a tilde symbol. Amino acid sequences were obtained from the UniProtKB database (Apweiler, Bairoch et al. 2004).

Collagen turnover is exceptionally slow. Age-related accumulation of D-aspartic acid through racemization indicates articular cartilage and skin collagen has a half-life of 117 years and 14.8 years respectively (Verzijl, DeGroot et al. 2000). It has been suggested that slow collagen turnover may result from initial accessibility to binding sites in supramolecular collagen assemblies (Welgus, Jeffrey et al. 1980), this is supported by recent computational and molecular modelling of collagen fibrils in rat tail tendon which identified restrictions when accessing the MMP collagenase cleavage site due to quasihexagonal packing and the presence of obstructive C-telopeptides (Orgel, Irving et al. 2006, Perumal, Antipova et al. 2008). However, accelerated collagen destruction is associated with up-regulation of MMP expression as observed in a number of pathologies, as displayed in degenerative arthritic disease characterised by articular cartilage destruction, the progressive matrix lysis observed in periodontal disease, and proteolysis of ECM constituents and basement membranes facilitating tumour cell metastasis (Mehraban, Kuo et al. 1994, Shah, Kumar et al. 1994, Sapna, Gokul et al. 2013).

## 1.4 The Matrix Metalloproteinase Family

The MMP family is comprised of zinc- and calcium-dependent multidomain secreted endopeptidases known to degrade multiple ECM components *in vitro*, these including laminin, proteoglycans, fibronectin, entactin, as well as fibrillar collagens (Galloway, Murphy et al. 1983, Muir and Manthorpe 1992, Sires, Griffin et al. 1993). MMP activity is, however, not restricted to ECM substrates. Studies have shown a number of cytokines, cell-surface adhesion receptors and growth factors undergo MMP-mediated cleavage. Activity towards interleukin-8, N-cadherin, syndecan-1 and hyaluronan receptor CD44 indicate that MMPs may have significant involvement in a number of key physiological processes, such as neutrophil activation, cell migration, proliferation and invasion (Van den Steen, Proost et al. 2000, Kajita, Itoh et al. 2001, Endo, Takino et al. 2003, Nakamura, Suenaga et al. 2004, Su, Blaine et al. 2008, Williams, Johnson et al. 2010).

Fibroblast interstitial collagenase MMP-1 (EC number 3.4.24.7) was the first MMP to be described following observations of dissolution of the bullfrog tadpole tail and collagenolysis during amphibian metamorphosis (Gross and Lapiere 1962, Gross and Nagai 1965) and is now considered the archetypal member of the family.

Classified as a subfamily of the metzincin superfamily of proteases, all MMP members have a conserved structural topology, sharing similar primary, secondary and tertiary structure. Typically, MMPs contain an N-terminal signal sequence rich in hydrophobic residues that targets the enzyme to the endoplasmic reticulum prior to secretion, a propeptide (PRO) domain to maintain latency, a $Zn^{2+}$ containing catalytic (CAT) domain and a variant length proline-rich flexible linker region connecting the C-terminal hemopexin (HPX) domain (Li, Brick et al. 1995, Jozic, Bourenkov et al. 2005, Iyer, Visse et al. 2006, Nagase, Visse et al. 2006). Subtle structural variations on this central modular design (Figure 1.4), typified by MMP-1, contributes to some of the observed diverse substrate specificities.

Although many MMPs are multifunctional and frequently share substrate specificities, they are further subcategorised into six groups based on sequence similarity and domain organisation. These are collagenases, gelatinases, stromelysins, matrilysins, MT-MMPs and a miscellaneous group of family members that have yet to be fully characterised. General sub-family features and characteristics are covered briefly below. Readers are directed to a number of reviews (Massova, Kotra et al. 1998, Bode and Maskos 2003, Nagase, Visse et al. 2006, Page-McCaw, Ewald et al. 2007) and references therein for further information.

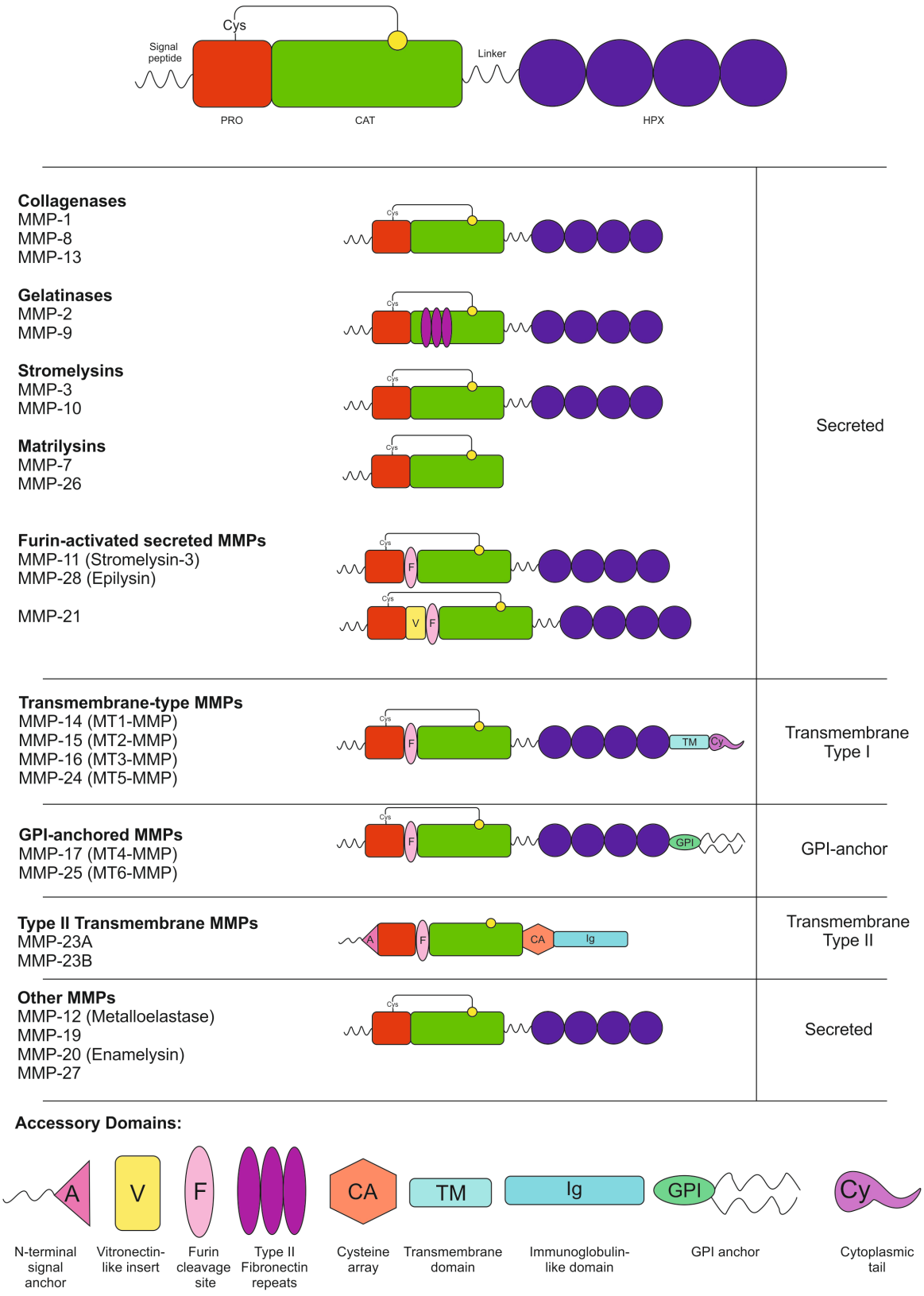



**Figure 1.4: Matrix metalloproteinase domain structure.** Structure and substrate specificity determine the sub-family classification. The five main groups are collagenases, gelatinases, stromelysins, matrilysins, membrane-bound MMPs. Other MMP members can not be classified due to non-typical features. The basic domain structure is characterised by the presence of a signal peptide, an inhibitory propeptide (PRO), a zinc-binding catalytic domain (CAT) and a hemopexin-like C-terminal domain (HPX) attached by a variable length linker. Addition of accessory domains to the basic topology enhance both structural and functional diversity (Adapted from Bourboulia and Stetler-Stevenson 2010).

### 1.4.1 Collagenases

MMP-1, -8 and -13 are interstitial collagenases secreted by multiple cell types in response to the stimulatory effects of growth factors and pro-inflammatory cytokines, primarily TNFα and IL-Iβ, upregulating gene expression (Borden, Solymar et al. 1996). These MMPs are principally responsible for degradation of fibrillar collagen (types I-III), which is imparted by cooperative action of the CAT and HPX domains (Murphy, Allan et al. 1992, Hirose, Patterson et al. 1993, Knauper, Cowell et al. 1997, Chung, Yoshida et al. 2004). Conserved residues Tyr210, Asp231 and Gly233 have also been described as a collagenase specificity determinant and are a unique feature of all collagenases from humans, rabbits and pigs (Sanchez-Lopez, Alexander et al. 1993, Freije, Diez-Itza et al. 1994).

Ubiquitously expressed, MMP-1 was the first vertebrate collagenase purified to homogeneity (from human skin fibroblasts) and cloned as cDNA (Goldberg, Wilhelm et al. 1986). MMP-1 is secreted in two zymogenic forms; a minor 57 kDa protein with N-linked glycosylation of Asn120 (Wilhelm, Eisen et al. 1986, Saarinen, Welgus et al. 1999) and a predominant unglycosylated protein of 52 kDa. The physiological role of MMP-1 glycosylation modification has yet to be elucidated. Cleavage of the PRO domain produces a mature enzyme composed of a CAT domain tethered to a HPX domain, a structure typical of both collagenases and stromelysins.

MMP-1 is generally undetectable in normal resting tissue, but is associated with a number of disease states. Constitutive expression accompanied by collagenolysis has been observed in aggressive and invasive human carcinomas (Templeton, Brown et al. 1990, Benbow, Schoenermark et al. 1999, Huntington, Shields et al. 2004, Eck, Cote et al. 2009) as well as the articular cartilage destruction connected with rheumatoid arthritis (Konttinen, Ainola et al. 1999, Yoshihara, Nakamura et al. 2000, Tetlow, Adlam et al. 2001).

MMP-8 is primarily synthesised and stored in specific granules of the polymorphonuclear leukocytes of maturing bone marrow, and was cloned from mRNA extracted from a patient with granulocytic leukaemia (Hasty, Pourmotabbed et al. 1990), hence the common name neutrophil collagenase. MMP-8 is released upon stimulation by chemical signals that trigger degranulation (Schettler, Thorn et al. 1991). Its presence in chondrogenic tissues has been reported (Cole, Chubinskaya et al. 1996) and work in a murine model showed that MMP-8 is expressed during embryogenesis (Giambernardi, Sakaguchi et al. 2001). MMP-8 exists as a highly-glycosylated 75 kDa

enzyme with an estimated carbohydrate content of 35% (Knauper, Kramer et al. 1990). It has 57% amino acid sequence homology to MMP-1 (Hasty, Pourmotabbed et al. 1990), and has a 15-fold preference for type I collagen over type III (Horwitz, Hance et al. 1977).

The third member of the collagenase sub-family, MMP-13, was identified in 1994 during a search for putative MMPs involved in breast carcinoma (Freije, Diez-Itza et al. 1994). Due to it's broad substrate specificity (Knauper, Cowell et al. 1997) and marked increase in expression in various invasive metastatic carcinomas (Johansson, Airola et al. 1997, Airola, Karonen et al. 1999), MMP-13 was considered a potent pathological collagenase. However, it has a substantial role in skeletal development; not only is MMP-13 expression restricted to mineralizing skeletal tissue during embryogenesis (Stahle-Backdahl, Sandstedt et al. 1997), but also *Mmp-13*$^{-/-}$ mice show significant defects in cartilage formation concomitant with collagen accumulation and delayed ossification (Inada, Wang et al. 2004).

Despite shared global structural homology between collagenase members there are significant differences in collagen processing capability (Table 1.1). Extensive differences in the depth and composition of collagenase specificity pockets within the active site are known to impact inhibitor selectivity (Lovejoy, Welch et al. 1999), therefore, it is not unreasonable to assume that subtle structural variations within the catalytic domain may also affect collagenolysis. In addition, significant differences in catalytic efficiency has also been observed as a result of substrate variation. For example, comparison of MMP-1 action on human and guinea pig type III collagen showed a 30-fold difference in $k_{cat}/K_m$ (Welgus, Jeffrey et al. 1981). Equally there is disparity within the α2(I) collagenase cleavage site sequence between these two species. Such variation likely contributes to the observed species-dependent substrate catalytic efficiency of human collagenases.

Table 1.1: Kinetic parameters of human collagenases[ab]

| Enzyme | Substrate | $K_m$ (μM) | $k_{cat}$ ($s^{-1}$ . $10^{3)}$ | $k_{cat}/K_m$ ($M^{-1}$ $s^{-1}$ x $10^{-3}$) | Ref. |
|---|---|---|---|---|---|
| MMP-1 | Type I collagen | 0.8 | 15 | 18 | |
| | Type II collagen | 2.1 | 0.28 | 0.13 | Welgus, Jeffrey, et al. 1981 |
| | Type III collagen | 1.4 | 160 | 69 | |
| MMP-8 | Type I collagen | 0.7 | 1.8 | 2.5 | |
| | Type II collagen | 1.1 | 0.65 | 0.59 | Hasty, Jeffrey et al. 1987 |
| | Type III collagen | 1.8 | 0.23 | 0.13 | |
| MMP-13 | Type I collagen | ND | ND | ND | |
| | Type II collagen | 2 | 6.3 | 3.9 | Mitchell et al. 1996 |
| | Type III collagen | ND | ND | ND | |

[a] Data obtained at 25 °C, pH 7.5
[b] ND: Not Determined
Adapted from Cancer Drug Discovery and Development: Matrix Metalloproteinase Inhibitors in Cancer Therapy
Edited by: Neil J. Clendeninn and Krzysztof Appelt © 2011 Humana Press Inc., Totowa, NJ

### 1.4.2 Gelatinases

Gelatinases (MMP-2 and -9) have the typical secreted MMP structure with the addition of a fourth domain distinguished by tandem triple fibronectin type II (F2) domains inserted within the CAT domain (Morgunova, Tuuttila et al. 1999). These enzymes readily digest gelatin, elastin, as well as type IV collagen, a major constituent of the basal lamina (Murphy, Cockett et al. 1991) and are grouped together by structural homology and enzymatic activity.

The MMP-2 gene was initially characterised and reported as a potential house-keeping gene due to an apparent lack of regulatory elements in the promoter region (Huhtala, Chow et al. 1990). This was supported by constitutive expression of the 72 kDa MMP-2 enzyme in a number of cell types and cancer cell lines (Qin, Moellinger et al. 1998, Roomi, Monterrey et al. 2009) as well as unaltered mRNA expression after treatment with MMP inducers IL-1 and TNα (Mackay, Ballin et al. 1992). However, mRNA production has been stimulated in specific tissues as seen with TGF-β treatment (Overall, Wrana et al. 1991, Marti, Lee et al. 1994) and it now appears that regulation of MMP-2 may be much more complex than previously anticipated.

Removal of the PRO domain proved to be another area of complexity for this particular gelatinase when it became apparent that activation was the result of cell-surface localization mediated by both membrane-bound MMP-14 and Tissue inhibitor of metalloproteinase-2 (TIMP-2), an endogenous MMP inhibitor (Itoh, Binner et al. 1995, Fernandez-catalan, Bode et al. 1998, Overall, Tam et al. 2000). In addition, a published crystal structure of the zymogen showed an interaction between a phenylalanine residue of the PRO domain and the third F2 domain (Morgunova, Tuuttila et al. 1999). Clearly this interaction must be disrupted, uncovering substrate-binding sites, thereby allowing full gelatin degrading properties to emerge (Banyai, Tordai et al. 1994, Murphy, Nguyen et al. 1994, Xu, Wang et al. 2004).

Despite recognition as a gelatinase, full-length MMP-2 has been reported as having the capacity to degrade fibrillar collagens (Konttinen, Ceponis et al. 1998) with a reported type I collagen affinity of 8.5 μM (Aimes and Quigley 1995). This is not a general feature of the gelatinase family, as MMP-9 lacks collagenolytic activity against collagen types I and II (Aimes and Quigley 1995, Konttinen, Ceponis et al. 1998) despite sharing F2 domain homology. The F2 domains of MMP-2 appear to direct gelatin-binding properties (Banyai, Tordai et al. 1994, Murphy, Nguyen et al. 1994, Xu, Wang et al. 2004), as well as type I collagen-binding activity (Tam, Moore et al. 2004), unlike the collagenases which utilise the C-terminal HPX domain. However, the HPX domain of MMP-2 also remains important because without it collagenolytic behaviour is lost (Steffensen, Wallon et al. 1995, Wallon and Overall 1997, Patterson, Atkinson et al. 2001, Tam, Moore et al. 2004). Therefore, MMP-2 requires both the presence of the HPX domain and F2 domains for collagenolysis, and operates a fundamentally different mode of action compared to the collagenases.

MMP-9 is distinctive: when found secreted in human lung fibroblasts the 92 kDa enzyme was found to have an additional flexible fifth domain with sequence homology

to a segment of the α2(V) collagen chain (Wilhelm, Collier et al. 1989). This domain is approximately 60 residues in length, is particularly rich in proline, serine and threonine and is heavily glycosylated. In fact, around 85% of the total sugar content is O-linked glycans attributed to this flexible linker domain (Mattu, Royle et al. 2000).

A combined small angle x-ray scattering (SAXS) and atomic force microscopy (AFM) study of full-length and C-terminally truncated MMP-9 investigated the inherent flexibility of the linker domain. Significant global conformational changes were reported highlighting the potential role of dynamic domain mobility for full-length enzyme-substrate interactions (Rosenblum, Cohen et al. 2007).

Homodimerisation, a feature not seen in other secreted MMPs, has also been reported for MMP-9. The asymmetric dimer is formed via non-covalent interactions between HPX domains (Olson, Bernardo et al. 2000, Cha, Kopetzki et al. 2002). Dimerization has been observed effecting both the autolytic activation mechanism (Olson, Bernardo et al. 2000) and MMP-9 mediated COS-1 cell migration (Dufour, Zucker et al. 2010), however regulation and function of dimerization has yet to be elucidated fully.

### 1.4.3 Stromelysins

MMPs -3, -10 and -11 constitute the stromelysin subfamily. They are secreted enzymes containing the simple structural arrangement seen in collagenases but lacking the ability to hydrolyse the fibrillar collagens (type I-III).

Comparison of MMP-3 and MMP-1 amino acid sequences demonstrate the close homology (~54%) between the members (Whitham, Murphy et al. 1986, Saus, Quinones et al. 1988). This, combined with the location of stromelysin and collagenase genes clustered on chromosome 11 (Formstone, Byrd et al. 1993), suggests a common evolutionary origin. However, despite the above homology, and presence of a collagen-binding hemopexin domain (Allan, Hembry et al. 1991), MMP-3 is unable to cleave type I collagen (Wilhelm, Collier et al. 1987).

MMP-3 was originally purified from human rheumatoid synovial cells (Okada, Nagase et al. 1986) displaying proteolytic activity against a broad range of ECM substrates, including fibronectin, laminin and type IV collagen (Galloway, Murphy et al. 1983, Chin, Murphy et al. 1985, Wilhelm, Collier et al. 1987, Okada, Konomi et al. 1989), as well as autoproteolytic activity that removes it's own HPX domain (Okada, Harris et al. 1988). Another feature, first described by observing MMP-3 proteolysis,

was it's ability to activate a number of MMPs *in vitro* by cleaving within the N-terminal PRO domain. It has activity against proMMP-9 (Ogata, Enghild et al. 1992) and -8 (Knauper, Wilhelm et al. 1993) as well as being an activator of proMMP-1 cleaving specifically at the Gln80-Phe81 scissile bond generating a mature MMP-1 enzyme with enhanced proteolytic activity (Murphy, Cockett et al. 1987, Suzuki, Enghild et al. 1990).

MMP-10 by contrast is much less well-characterised but is anticipated to have similar properties to MMP-3 due to significant homology between both the catalytic and HPX domains, 86% and 74% respectively (Bertini, Calderone et al. 2004). As an MMP activator, this is certainly true of MMP-10 with activity observed against proMMPs -1, -8 and -9 (Windsor, Grenett et al. 1993, Knauper, Murphy et al. 1996, Nakamura, Fujii et al. 1998). Therefore, both MMP-3 and -10 can indirectly affect collagen degradation by participating in an activation cascade of latent collagenases.

Although originally described as the third member of the stromelysin subfamily, due to overexpression in the stromal cells surrounding invasive breast neoplastic cells (Basset, Bellocq et al. 1990) and structural similarities, MMP-11 is arguably more suited to reclassification as it has a number of additional features. These include a furin activation sequence that is processed by a transmembrane serine protease in the *trans*-golgi network (Pei and Weiss 1995); an inability to be activated by 4-aminophenylmercuric acetate (APMA) (Santavicca, Noel et al. 1996), an organomercurial that typically induces activation in other secreted MMPs; and an unusual alanine substitution in place of a highly conserved proline residue within the CAT domain (Noel, Santavicca et al. 1995). This structural alteration has been suggested to affect enzyme activity and hence the restricted substrate profile associated with MMP-11. This profile does not include the classic ECM constituents normally associated with MMPs (Murphy, Segain et al. 1993, Pei, Majmudar et al. 1994). A recent study has demonstrated MMP-11 activity against cellular substrate galectin-1 (Kleifeld, Doucet et al. 2010), which has been associated with increased tumour growth and spontaneous metastasis (Banh, Zhang et al. 2011).

### 1.4.4 Matrilysins

MMP-7 and -26 are known as matrilysins, MMPs containing the minimal domain requirements necessary for secretion, latency and activity. As such they are functionally distinctive lacking the characteristic HPX domain that is often implicated in substrate selection and unsurprisingly are unable to cleave fibrillar collagens (Miyazaki, Hattori et

al. 1990, Uria and Lopez-Otin 2000). They also possess a threonine residue adjacent to the zinc-binding site which appears to be a specific feature of all known matrilysins from human, rat, mouse and cat (Uria and Lopez-Otin 2000), the function of which, is unknown.

The cDNA encoding MMP-7 displays amino acid sequence homology of ~40 % with MMP-3 and MMP-1 (Muller, Quantin et al. 1988). The secreted protein was isolated from a rectal carcinoma line (Miyazaki, Hattori et al. 1990) and shows activity against a broad range of ECM glycoproteins including fibronectin, aggrecan and entactin expression (Woessner and Taplin 1988, Fosang, Neame et al. 1992, Sires, Griffin et al. 1993).

MMP-26 has the same structural topology as MMP-7 but in contrast is more closely related to the metalloelastase MMP-12 (de Coignac, Elson et al. 2000). The enzyme appears to have restricted expression in both uterine and placental tissues with the original cDNA cloned from human placental and endometrial tumour lines (Park, Ni et al. 2000, Uria and Lopez-Otin 2000) and therefore most likely has a distinctive role in the menstrual cycle and other reproductive processes.

### 1.4.5 MT-MMPs

Six membrane-bound MMPs have been reported and are defined by possession of a C-terminal extension that is involved in cell membrane insertion. They are subdivided into two groups; MMP-14, -15, -16 and -24 are classed as Type I transmembrane (TM) proteins and utilise a single pass cytoplasmic intracellular domain; the second group, comprised of MMP-17 and -25, have a glycosylphosphatidylinositol (GPI) anchor (Itoh, Kajita et al. 1999, Kojima, Itoh et al. 2000). Phosphorylation (Nyalendo, Michaud et al. 2007) and binding of regulatory proteins like gC1qR (Rozanov, Ghebrehiwet et al. 2002) via the cytoplasmic tail has been reported, validating suggestions of roles in intracellular signalling, localisation and trafficking which may subsequently impact cell migration (Gingras, Bousquet-Gagnon et al. 2001).

The basic tetrapeptide RRKR/RRRR sequence within the PRO domain (Takino, Sato et al. 1995, Puente, Pendas et al. 1996) is common to all membrane-bound MMPs as well as the secreted MMPs -11, -21, -23, -28. Recognition of this motif by furin was considered sufficient to ensure these proteins were secreted from the golgi in an activated state (Pei and Weiss 1995, Sato, Kinoshita et al. 1996, Yana and Weiss

2000). However, a recent study demonstrated enzyme maturation required multi-proteolytic steps (Golubkov, Cieplak et al. 2010).

Similar to the stromelysins, certain members of this subfamily have been identified as major MMP activators. MMP-14 was initially recognised as a key player in proMMP-2 activation (Will, Atkinson et al. 1996, Itoh, Takamura et al. 2001) and has since been proposed to be an activator of proMMP-13 (Knauper, Will et al. 1996, Knauper, Bailey et al. 2002), although the mechanism has yet to be fully elucidated. MMP-15, -16, -24 and -25 have also been reported to process proMMP-2 (Pei 1999, Morrison, Butler et al. 2001, Nie and Pei 2003, Zhao, Bernardo et al. 2004, Morrison, Butler et al. 2009). This action will most certainly amplify the proteolytic cascade directly at the cell surface, i.e. in the pericellular environment.

In addition to degrading ECM constituents such as type I collagen (d'Ortho, Will et al. 1997, Ohuchi, Imai et al. 1997) these enzymes can target a number of cell-surface proteins, such as transglutaminase, CD44 and syndecan-1 (Belkin, Akimov et al. 2001, Kajita, Itoh et al. 2001, Endo, Takino et al. 2003). Presumably contributing to complex cell-signalling pathways.

### 1.4.6 Others

A further seven MMPs exist that have not been classified in the other categories. These include MMP-12, MMP-20 and MMP-23, among others which demonstrate the great structural and functional diversity displayed by this family of metalloproteases.

MMP-12 is a metalloelastase originally found expressed in the alveolar macrophages of cigarette smokers (Shapiro, Kobayashi et al. 1993). It is particularly effective against elastin and is therefore important in mediating elastin turnover. MMP-12 is an abundant elastase found in lung and arterial wall, consequently it has been associated with both emphysema and atherscleros is (Liang, Liu et al. 2006, Yamada, Wang et al. 2008, Churg, Zhou et al. 2012). Like many MMPs, the gene is not transcribed in normal resting tissues. Gene expression has been identified in both the macrophage and stromal cells of the placenta. Analysis of the amino acid sequence indicates homology with collagenase MMP-1 and stromelysin MMP-3 (Belaaouaj, Shipley et al. 1995).

MMP-19, first isolated from a patient with rheumatoid arthritis (Kolb, Mauch et al. 1997), possesses the typical domain organisation observed in secreted MMPs. However, it also exhibits two distinctive features, a threonine-rich sequence at the C-

terminus and an oligoglutamate stretch in the linker region (Sedlacek, Mauch et al. 1998). Expression has been identified in a wide distribution of tissues including placenta, lung, pancreas, ovary, spleen and intestine (Pendas, Knauper et al. 1997), suggesting a possible role in normal tissue homeostasis.

MMP-20, also known as enamelysin, was originally identified from a porcine enamel organ-specific cDNA library (Bartlett, Simmer et al. 1996). MMP-20 is tooth-specific and digests amelogenin (Ryu, Fincham et al. 1999), the major protein constituent in the enamel matrix of teeth. Defects in enamel bio-mineralization are the principal characteristic of amelogenin imperfecta and mutations of MMP-20 have been implicated in this disorder (Caterina, Skobe et al. 2002, Kim, Simmer et al. 2005).

MMP-21 is the human orthologue of XMMP, a matrix metalloproteinase transiently expressed at the initiation of gastrulation of *Xenopus laevis* (Yang, Murray et al. 1997). MMP-21 was isolated from a human placenta cDNA library (Ahokas, Lohi et al. 2002) and has been seen expressed during human embryogenesis as shown by it's presence in fetal kidney, brain and liver tissues (Marchenko, Marchenko et al. 2003). Structurally similar to most MMPs, MMP-21 has a furin activation sequence, an unconserved proline-rich insertion between the PRO and CAT domain that is seen in both *X. laevis* and *Cynops pyrrhogaster* (Japanese Fire-bellied newt) orthologues (Suzuki, Tadano et al. 2001), while lacking a linker region connecting the HPX and CAT domains (Marchenko, Marchenko et al. 2003).

A particularly unusual member is MMP-23. Originally isolated during characterisation of the terminal end of chromosome 1, MMP-23 is encoded by two genes located within two identical tandemly-linked genomic regions (Gururajan, Grenet et al. 1998, Gururajan, Lahti et al. 1998). MMP-23 was subsequently cloned from an ovarian cDNA library and appears to be expressed predominantly in reproductive tissues of the ovary, testis and prostate (Velasco, Pendas et al. 1999). MMP-23 has some unique features that include an N-terminal transmembrane signal sequence (Pei, Kang et al. 2000), a shortened PRO domain lacking the typical 'cysteine switch' motif and a truncated C-terminal domain with no homology to the HPX domain. This has been "replaced" with a novel cysteine array motif (Cys-$X_{(6)}$-Cys-$X_{(8)}$-Cys-$X_{(10)}$-Cys-$X_{(3)}$-Cys-$X_{(2)}$-Cys), and an Ig-like domain (Pei 1999). These features aside, MMP-23 has been classified as a matrix metalloproteinase due to the homology of it's CAT domain and possession of a furin recognition motif (Pei, Kang et al. 2000).

MMP-27 is potentially the last uncharacterised human MMP. Initially it was cloned from the sclera of *Tupaia belangeri* (Common tree shrew) and with sequence data

deposited in the GenBank sequence database (accession number AF281673) by Guggenheim and Frost in 2000 (unpublished). The human orthologue was identified during a large-scale bioinformatic search of novel human secreted and transmembrane proteins (Clark, Gurney et al. 2003). It's substrate specificity and physiological relevance have yet to be determined.

MMP-28 (epilysin) is expressed in a wide range of healthy tissues with constitutive expression observed in the testis, lung epithelium, invertebral disc, heart, colon and intestine (Lohi, Wilson et al. 2001, Marchenko and Strongin 2001, Gruber, Ingram et al. 2009, Manicone, Birkland et al. 2009). Interestingly, mRNA levels for MMP-28 appear to be unaffected by a number of pro-inflammatory cytokines (Saarialho-Kere, Pentland et al. 1994, Klawitter, Quero et al. 2011). This is in contrast to a number of other MMPs which are upregulated in the presence of TNF-α, IL-1α and IL-1β (Hanemaaijer, Koolwijk et al. 1993, Makela, Salo et al. 1998, Solomon, Li et al. 2000). Therefore, it appears that MMP-28 functions in normal tissue homeostasis.

Phylogenetic analysis indicates it is closely related to MMP-19 (~40%) but in contrast it possesses the furin activation motif and an atypical promoter, noted for absence of the regulatory TATA box (Illman, Keski-Oja et al. 2001, Marchenko and Strongin 2001).

## 1.5 Matrix Metalloproteinase Domain Structure

### 1.5.1 Inhibitory Domain

Enzyme maturation and catalytic activity acquisition occur as a direct result of propeptide cleavage. Our understanding regarding CAT and PRO domain interactions and the mechanism of auto-inhibition has been enhanced by published structures of full length proMMPs -1 (Jozic, Bourenkov et al. 2005) and -2 (Morgunova, Tuuttila et al. 1999), and C-terminally truncated proMMPs -3 (Becker, Marcy et al. 1995) and -9 (Elkins, Ho et al. 2002), which are virtually superimposable over the main-chain atoms (Figure 1.5).

Essentially, the propeptide is a small globular domain of ~ 80 residues that form three α helices and connecting loops. This is stabilised by hydrogen bonds between the helices and formation of a hydrophobic core, in proMMP-1 this is supported by the side-chain interaction between Asp61 and Gln49. Latency is maintained through steric blockage of the active site, excluding water as a potential fourth ligand for the zinc-ion,

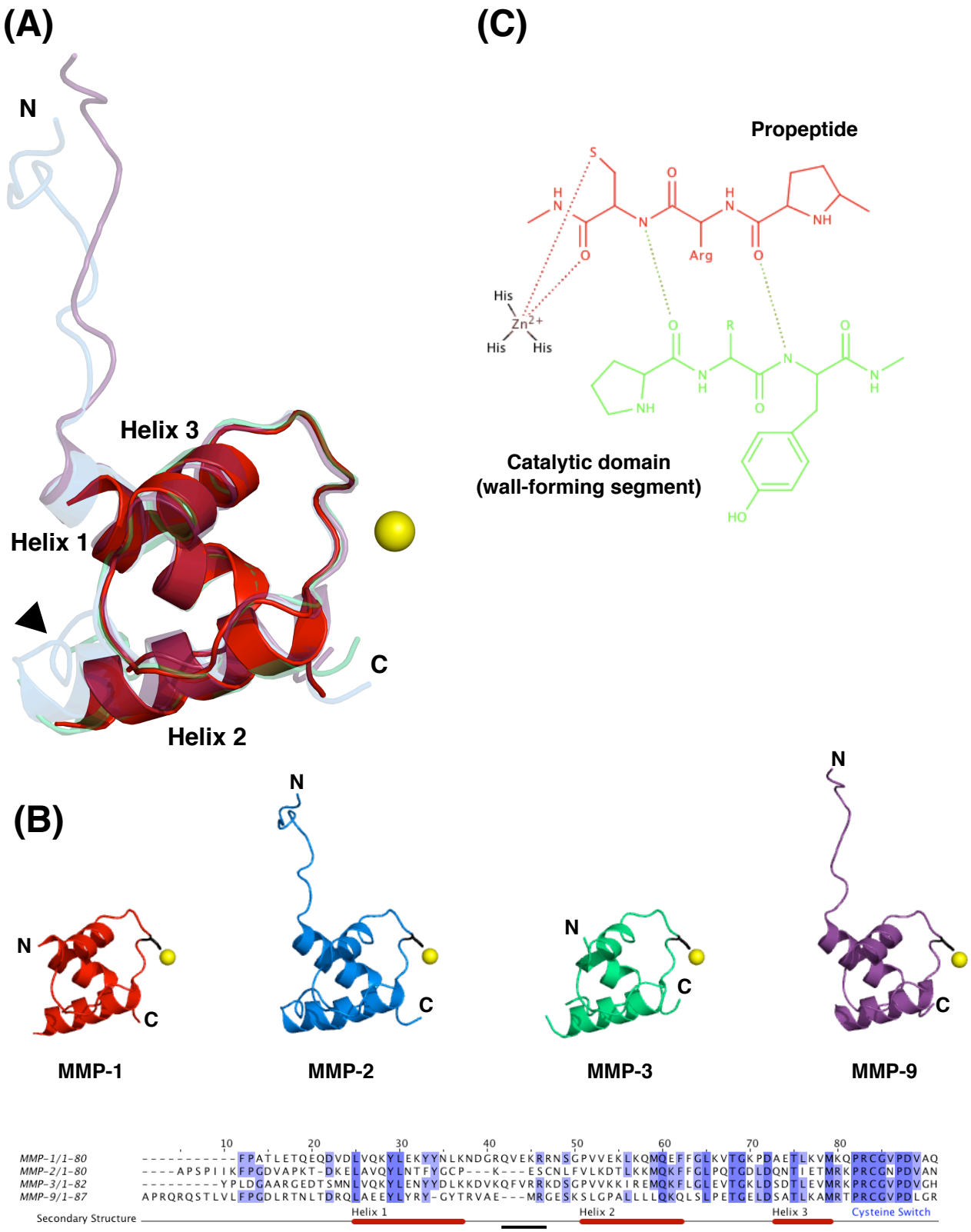



**Figure 1.5: Structure of the inhibitory propeptide.** (**A**) Superimposition of the propeptide of MMP-1 (1SU3.pdb), MMP-2 (1CK7.pdb), MMP-3 (1SLM.pdb) and MMP-9 (1L6J.pdb). Structures were aligned against proMMP-1 to minimize the RMSD interatomic distance difference on main-chain atoms (0.47-0.78 Å) using Pymol Molecular Graphics System (version 1.5.0.4.), Schrödinger, LLC. The zinc ion is represented as a *yellow* sphere. Protease sensitive loop 1 is indicated by the filled triangle. (**B**) Amino acid sequence alignment and individual views of relevant MMP propeptides. Alignment is coloured with increasing percentage identity (75-100%). 'Bait region' motif of proMMP-1 is underlined. Invariant cysteine residue of the 'cysteine switch' motif is shown in *black* stick form coordinating the catalytic zinc ion. (**C**) Schematic representation of interactions between an MMP and propeptide. The sulfhydryl group of the unpaired cysteine residue forms an intramolecular tetrahedral complex with the zinc-ion binding histidine triad of the catalytic domain. Propeptide is in red and catalytic domain is in green. Image rendered using MarvinSketch (version 5.10.3.), ChemAxon.

and occluding substrate entry. This interaction is achieved by the presence of an invariant cysteine residue found at position 73 within the highly conserved inactivation motif P71-R-C-G-X-P-D77 known as the 'cysteine switch' (Springman, Angleton et al. 1990, Van Wart and Birkedal-Hansen 1990). All known MMPs share this signature feature, with the exception of MMP-23. Mutations in the cysteine switch consensus sequence have been shown to significantly reduce latency. For example, replacement of the critical cysteine residue with other potential zinc-binding residues leads to complete loss of inhibitory activity (Park, Matrisian et al. 1991).

Other residues contribute to PRO domain stabilisation. Tyr18 and Leu19 in helix 1 are conserved in a number of MMPs and MMP-3 mutagenesis experiments involving these residues demonstrates their role as N-terminal stabilisation motifs, as alanine substitution resulted in increased autolytic activity (Freimark, Feeser et al. 1994). Furthermore, a number of hydrogen bonds link the PRO and CAT domains (Figure 1.6). Solved structures of MMPs equipped with accessory domains also display non-covalent interactions between the domains. proMMP-1 has residues located within blade I of the HPX domain that appear to support the propeptide thus promoting a compact zymogenic state (Jozic, Bourenkov et al. 2005). In the case of the gelatinases, a phenylalanine is buried in a hydrophobic pocket on the surface of the third F2 repeat (Morgunova, Tuuttila et al. 1999, Elkins, Ho et al. 2002). In MMP-2 this is held in place by a salt bridge between Asp40 and Arg368 and hydrogen bonding between Ile35 and Gly367.

Loop regions are frequently missing in the electron density as observed in the proMMPs -1, -3 and -9 crystal structures. Protease-sensitive "bait regions" like that of proMMP-1 are known to reside within solvent exposed loop 1 (Nagase, Enghild et al. 1990) and cleavage within this region is thought to destabilise helix-helix interactions and the overall propeptide fold allowing further proteolysis. MMP-2, in contrast, has a unique disulphide bond between Cys60-Cys65 (Morgunova, Tuuttila et al. 1999) presumably conferring rigidity to loop 1, and which may feature in the specific MT-MMP/TIMP-2 activation mechanism.

### 1.5.2 Catalytic Domain

The MMP CAT domain is well documented with over 200 structures solved and available in the Protein Data Bank (Table 1.2). Roughly spherical in shape and ~40 Å in diameter (Tallant, Marrero et al. 2010), the CAT domain is composed of ~165

residues. The tertiary structure of this particular domain consists of 3 α-helices and 5 twisted β-strands (Figure 1.6) with a minimum of two calcium-binding sites and an additional two zinc-binding sites, one of which is critical for catalysis. The domain is asymmetrically bisected by the shallow substrate-binding cleft forming an N-terminal upper subdomain, which terminates at Gly225, followed by a much smaller lower C-terminal subdomain (Bode, Fernandez-Catalan et al. 1999). This configuration is almost identical in all currently known MMP structures.

Removal of the PRO domain is required for enzyme maturation. The method of activation is known to significantly impact the proteolytic activity status of the MMP CAT domain. Catalytic efficiency is highly dependent on the specific N-terminal residue generated during the activation process. MMP-1 activity is known to be enhanced by the action of MMP-3 (Murphy, Cockett et al. 1987), this is compared to other activation methods which lead to only ~14-25% of maximal activity. This observed superactivation occurs as a result of Phe81 occupying the N-terminal position (Nagase, Suzuki et al. 1992). In the case of MMP-8, a stromelysin-specific cleavage after Gly78 is necessary to generate the most effective enzyme (Knauper, Wilhelm et al. 1993, Knauper, Murphy et al. 1996). This results in a salt bridge between the ammonium group of the N-terminal Phe79 and the side chain carboxylate group of Asp232 (Reinemer, Grams et al. 1994, Iyer, Visse et al. 2006), which is thought to stabilise the 'Met-turn' positioned at the base of the active site.

Hydrolytic activity of mature collagenases occurs within the 5Å wide active site cleft, which extends horizontally (Figure 1.6C) with generation of a β-turn containing a strictly conserved methionine eight residues after the consensus sequence HExxHxxGxxH (Van Wart and Birkedal-Hansen 1990). This characteristic 'Met-turn' then forms a hydrophobic base beneath the catalytic zinc ion (Bode, Gomis-Ruth et al. 1993). The zinc ion is coordinated by the imidazole Nε atoms of the three invariant histidine residues (Figure 1.6) and $H_2O$ as a fourth ligand (Li, Brick et al. 1995, Iyer, Visse et al. 2006). Mutagenesis of any of the above glutamate and histidine residues has been demonstrated to ablate catalytic activity (Sanchez-Lopez, Nicholson et al. 1988, Windsor, Bodden et al. 1994, Roeb, Behrmann et al. 2000, Steele, El-Kabbani et al. 2000, Ozdemir, Hart et al. 2005).

Substrate binding involves the peptide bond carbonyl group coordinating with the active site $Zn^{2+}$ ion. Positioning of the carboxylate group of the essential catalytic glutamate residue deprotonates zinc-bound $H_2O$ permitting a water-based nucleophilic attack on the substrate scissile bond (Pelmenschikov and Siegbahn 2002).

Table 1.2: Experimentally derived Protein Data Bank accession codes for all MMP structures containing at least the catalytic domain (www.pdb.org; closing date: October 16, 2012).

| Subfamily | MMP | PDB Codes |
|---|---|---|
| Collagenase | MMP-1 | 966C 1AYK 1CGE 1CGF 1CGL 1FBL 1HFC 1MNC 2AYK 2TCL 3AYK 4AYK 1SU3 2J0T 2CLT 3SHI 4AUO |
| | MMP-8 | 1JH1 1JJ9 1KBC 1MMB 1MNC 1ZVX 2OY4 1ZS0 1JAN 1A85 1A86 1BZS 1I73 1I76 1JAO 1JAP 1JAQ 1ZP5 2OY2 1JH1 3DPE 3DPF 3DNG 1ZS0 |
| | MMP-13 | 456C 830C 1CXV 1EUB 1FLS 1FM1 2PJT 2OZR 2OW9 2E2D 2D1N 1ZTQ 1YOU 1XUR 1XUD 1XUC 3ELM 2YIG 317G 3I7I 3KEC 3KEJ 3KEK 3KRY 3LJZ 3O2X 3TVC 3ZXH 4A7B |
| Gelatinase | MMP-2 | 1CK7 1HOV 1QIB 1GXD 1EAK 3AYU |
| | MMP-9 | 1GKC 1GKD 1l6J 2OW0 2OW1 2OW2 2OVZ 2OVX |
| Stromelysin | MMP-3 | 1B3D 1B8Y 1BIW 1BM6 1BQO 1C3I 1CAQ 1CIZ 1CQR 1D7X 1D8F 1G49 1G4K 1OO9 1QIA 1QIC 1SLM 1SLN 1UEA 1UMS 1UMT 1USN 2D1O 2JNP 2JT5 2JT6 2USN 3OHL 3OHO 3USN 1C8T 1D8M 1D5J 1G05 1HY7 1HFS 2USN 1B8Y |
| | MMP-10 | 1Q3A 3V96 |
| | MMP-11 | 1HV5 |
| | MMP-12 | 1JIZ 1JK3 2Z2D 2K2G 3BA0 2OXU 2OXZ 2OXW 2HU6 1Z3J 1YCM 1Y93 1RMZ 1OS9 1OS2 2POJ 1UTZ 1UTT 1ROS 2W0D 3EHX 3EHY 1ZRG 2JPV 3F15 3F16 3F17 3F18 3F19 3FLA 2K9C 2JXY 2KRJ 2OXU 2OXW 2OXZ 2POJ 2W0D 2WO8 2WO9 2WOA 2Z2D 3BA0 3EHX 3EHY 3F15 3F16 3F17 SF18 3F19 3F1A 3LIK 3LIL 3LIR 3LJG 3LK8 3N2U 3N2V 3NX7 3RTS 3RTT 3TS4 3TSK 4EFS 4GUY 3LKA 3TT4 |
| Matrilysin | MMP-7 | 1MMP 1MMQ 1MMR 2DDY 2Y6C 2Y6D |
| Membrane-bound | MMP-14 | 1BQQ 1BUV 3MA2 |
| | MMP-16 | 1RM8 |
| Other | MMP-20 | 2JSD |

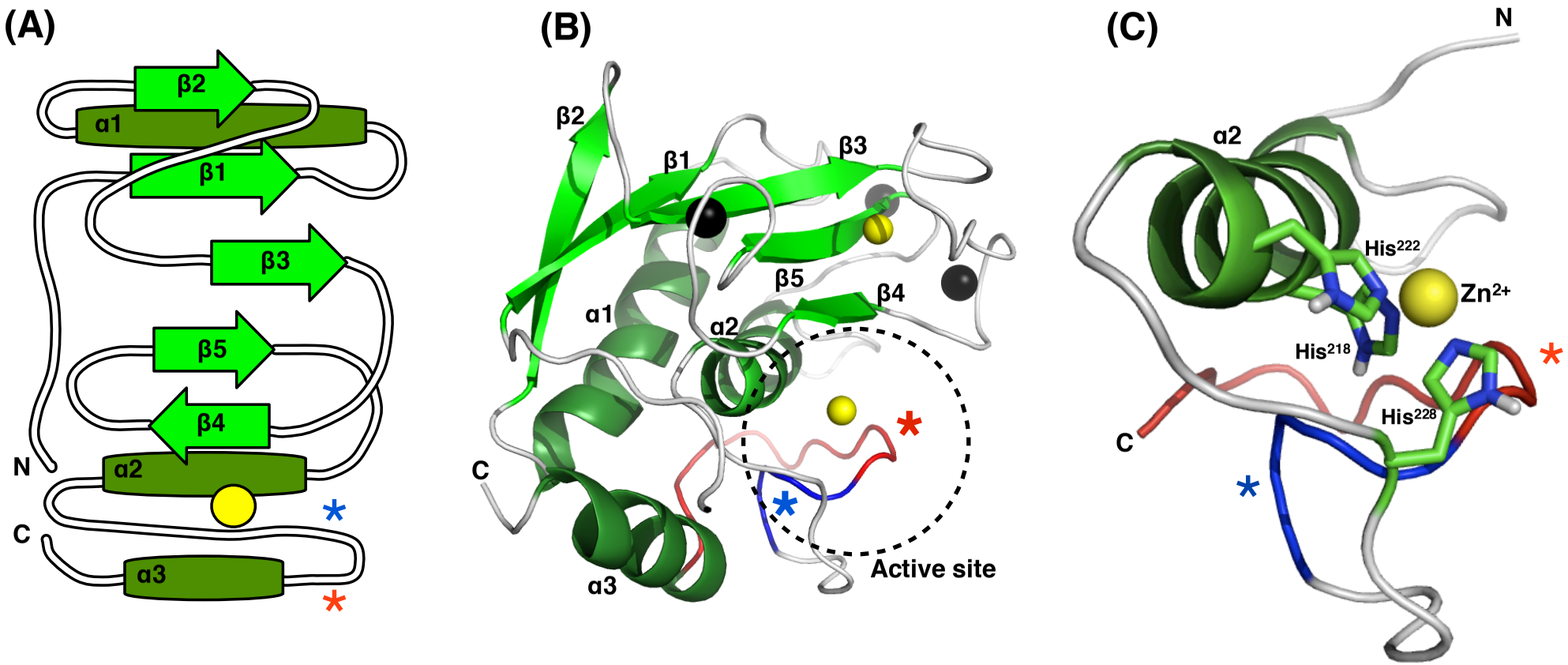



**Figure 1.6: Catalytic domain of MMP-1.** Zinc ions are depicted as *yellow* spheres and calcium ions are coloured *black*. Secondary structure elements, α helices and β strands, are coloured dark and light *green* respectively. Location of the 'Met-turn' is indicated by the *blue* star and the specificity loop is highlighted by the *red* star. (**A**) Schematic illustration of the MMP catalytic domain topology. For clarity only the catalytic zinc ion is shown. (**B**) Structure of MMP-1 (1CGE.pdb). The dashed circle highlights the active site, which is bordered by β4, α2 and a stretch of random coil adjacent to the N terminus of α3. (**C**) The active site cleft. The zinc-coordinating histidine triad is shown in stick form. The 'Met-turn' and specificity loop are coloured *blue* and *red* respectively.

Substrate catalysis results from a number of interactions located within the zinc-binding region directly responsible for cleavage of the peptide scissile bond. These contributions include substrate main chain binding within the active site cleft and specificity sub-site pockets that accommodate substrate side chain residues and mediate substrate preference. Additional contributions stem from potential exosites in the accessory domains that may act as secondary binding sites and/or manipulate the substrate prior to hydrolysis.

### 1.5.2.1 Specificity Sub-site Pockets

Several recognition pockets are distributed along the active site cleft, from S3-S3' (Babine and Bender 1997), mediating interactions with equivalent substrate peptide substituents (Figure 1.7). The S1' sub-site specificity pocket is the prime influencer of substrate specificity and compatibility within the CAT domain. The MMP-1 S1' pocket invaginates to the right of the active site zinc ion and is shaped by His218, Leu235 of the 'Met-turn' and the Pro238-Tyr240 wall forming segment (Bode, Fernandez-Catalan et al. 1999, Iyer, Visse et al. 2006). However, the presence of Arg214 tends to exclude entry of large side chains into the solvent filled pocket, thereby determining substrate preference (Bode, Fernandez-Catalan et al. 1999). This pocket diverges in shape and depth among MMPs as a result of leucine, arginine and tyrosine residues present at positions analogous to Arg214.

MMPs can be grouped dependent on S1' pocket depth (Figure 1.7). Shallow and intermediate pockets define MMP-1 and -7, and MMP-2 and -8 respectively. In contrast, MMP-3, -14 and -20 are distinguished by deep tube-like cavities (Yamamoto, Tsujishita et al. 1998, Lovejoy, Welch et al. 1999, Moy, Chanda et al. 1999, Pavlovsky, Williams et al. 1999, Park, Jin et al. 2003, Turk, Lee et al. 2006). These deep pocket MMPs can accommodate residues with much larger side chains (Mucha, Cuniasse et al. 1998). However, the CAT domain is far from rigid. Increased flexibility around the active site cleft has been reported for synthetic inhibitor-bound MMP-1, suggestive of either an induced-fit mechanism (Moy, Chanda et al. 1999) or conformational selection. This is supported by further observations using RS-104966-bound MMP, which displayed enhanced mobility in the backbone around Arg214 to accommodate the diphenylether backbone. Despite these structural observations, in the case of synthetic oligopeptides, mutations at the P1' position are rarely tolerated by MMP-1 (Netzel-Arnett, Fields et al. 1991), which prefers Ile/Leu at this position.

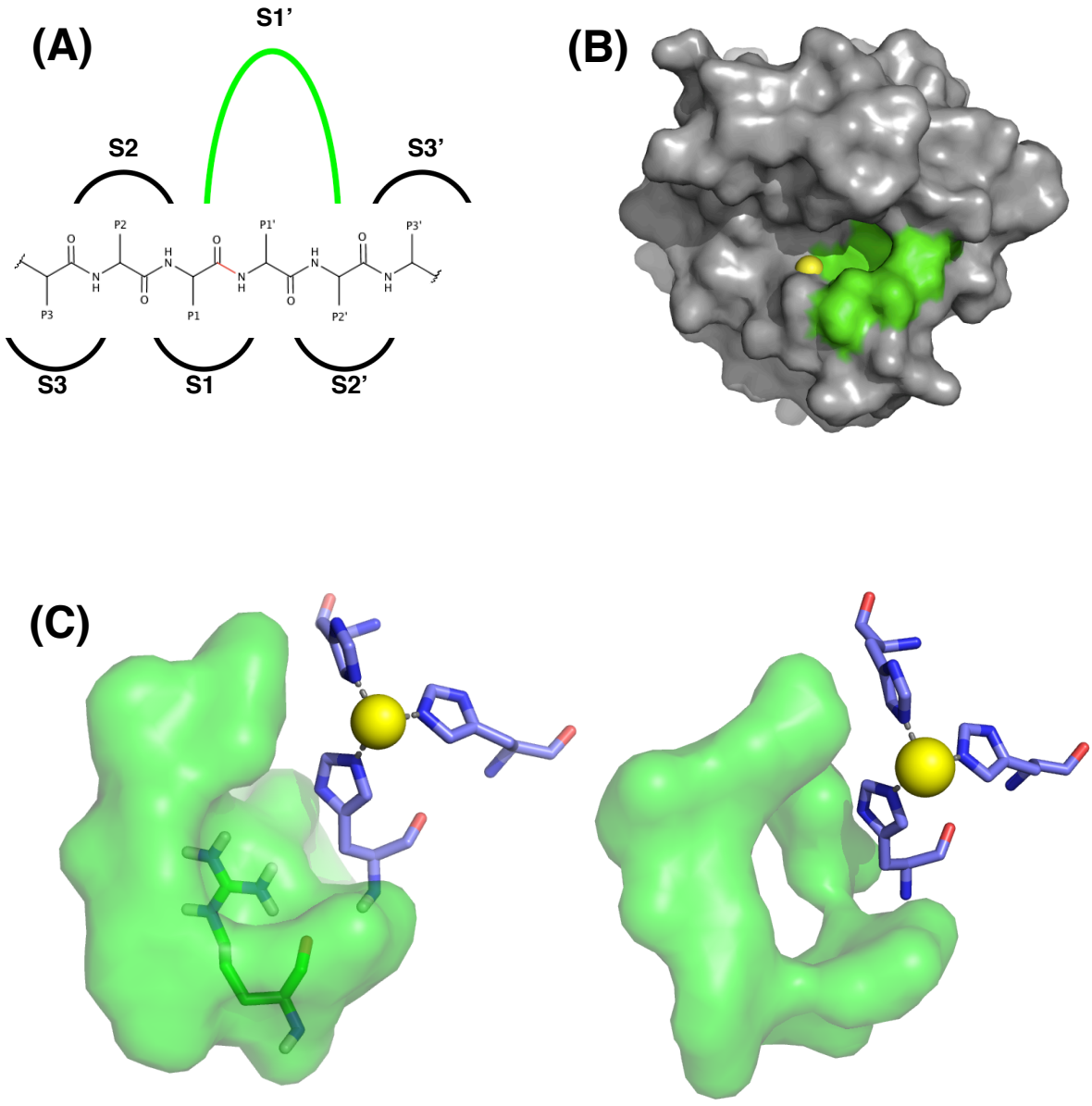



**Figure 1.7: The S1' specificity pocket.** (**A**) Schematic of a general peptide substrate bound within the active site of a protease. P*n* and P*n*' (*n* = 1, 2, 3) refer to substrate residues and unprimed S*n* and primed S*n*' refer to the cognate protease binding site. The MMP S1' (coloured *green*) site is the most variable binding pocket. (**B**) Surface view of MMP-1 (1HFC.pdb) in standard orientation. The zinc ion is depicted as a *yellow* sphere. The S1' pocket is located to the left side of the catalytic zinc residue and is defined by Arg214, His218, Leu235 and the wall forming segment Pro238-Ser239-Tyr240 (Bode, Fernandez-Catalan et al. 1999). (**C**) Variable S1' specificity pockets. MMP-1 (*left*) is occluded by Arg214 (*green sticks)* while MMP-3 (*right*) has a leucine at the equivalent position leading to an open channel.

### 1.5.3 Hemopexin Domain

The HPX domain is tethered to the CAT domain via a variable length linker. The linker region in MMP-1 is 16 residues long (Jozic, Bourenkov et al. 2005) and therefore

the HPX domain is within close proximity of the C-terminal helix of the CAT domain. In contrast, MMP-9 and -15 have much longer linkers, exceeding 60 residues in length (Wilhelm, Collier et al. 1989, Will and Hinzmann 1995).

For MMPs that possess the HPX domain, the topology is generally equivalent for sharing homology with rabbit serum hemopexin, a haem-transporting plasma glycoprotein (Faber, Groom et al. 1995). Comparison of resolved structures (MMP-1, -2, -9, -12, -13, -14) show a C-terminal domain of ~200 residues, forming a four blade β-propeller arrangement displaying pseudo four-fold symmetry that is stabilised by an intradomain disulphide bond between cysteine residues in blades I and IV (Gohlke, Gomis-Ruth et al. 1996, Gomis-Ruth, Gohlke et al. 1996, Morgunova, Tuuttila et al. 1999, Cha, Kopetzki et al. 2002, Jozic, Bourenkov et al. 2005, Iyer, Visse et al. 2006, Bertini, Calderone et al. 2008, Tochowicz, Goettig et al. 2011). Disruption of this bond by mutagenesis was reported to abolish collagenase activity by MMP-1 (Windsor, Birkedal-Hansen et al. 1991) and significantly impact collagenolysis by MMP-8 with alanine substitutions reducing activity against type I collagen by 62% (Hirose, Patterson et al. 1993).

A solvent-accessible funnel-shaped tunnel is positioned at the centre of the HPX domain and is shaped by the first strand of each β sheet. In MMP-1, the carbonyl oxygen atoms of Asp266, Asp359, Asp408 and Glu310 coordinate a central structural calcium ion (Iyer, Visse et al. 2006). MMP-2 possesses a similar acidic patch at the tunnel entrance, which is formed by four aspartate residues (Gohlke, Gomis-Ruth et al. 1996). Although calcium is typically found within the channel, sodium and chloride ions have also been observed (Gohlke, Gomis-Ruth et al. 1996, Gomis-Ruth, Gohlke et al. 1996, Tochowicz, Goettig et al. 2011). Whether these extraneous ions are an artifact of the crystallisation process, or have some physiological relevance, has yet to be established.

A single turn of α-helix and a four-stranded anti-parallel β sheet connected by long loops, constitute each blade (Figure 1.8). The majority of the surface contours are formed by these loop regions. Loop diversity, particularly associated with the outermost β-strand of each blade, is proposed to directly impart MMP-specific recognition properties. For example, collagen breakdown by MMP-1, cell-surface activation of proMMP-2 mediated by TIMP-2, and dimerisation of both MMP-9 and -14 is known to occur via the C-terminal HPX domain. MMP-9 forms asymmetric homodimers through reduction-sensitive bonding between loops in blade IV (Cha, Kopetzki et al. 2002) while membrane-bound MMP-14 forms symmetrical homodimers mediated

primarily by short loops in blade III (Tochowicz, Goettig et al. 2011). MMP-2 is localised at the cell-surface by blade III/IV interactions with intermediary TIMP-2 (Morgunova, Tuuttila et al. 2002). Conversely, MMP-1 is apparently reliant on determinants within blade I for collagen-substrate recognition (Lauer-Fields, Chalmers et al. 2009).

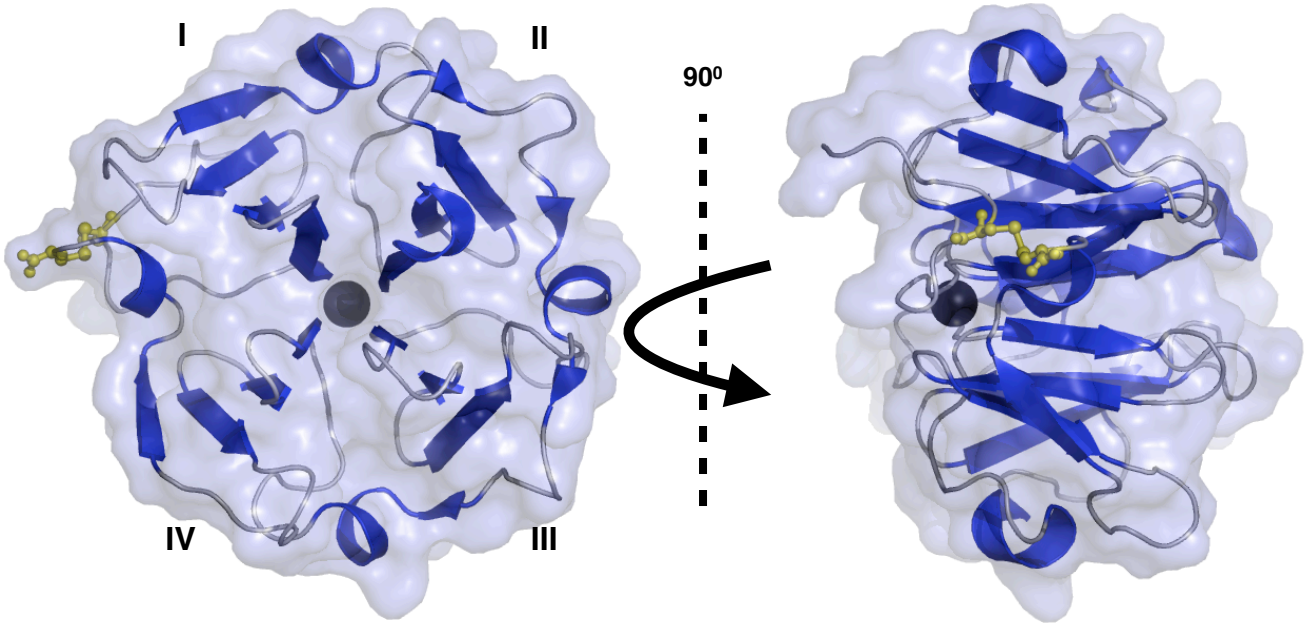



**Figure 1.8: MMP-1 hemopexin domain.** Structure of the C-terminal domain (~45 Å at the longest axis) showing semitransparent surface and ribbon representations in the same orientation. Blades I-IV are labelled clockwise. The calcium ion, depicted as a *black* sphere, is positioned at the base of the propeller channel. The ball and stick representation of the disulphide bond between blade I and IV is coloured *yellow*.

Although lacking inherent catalytic activity, the HPX domain of MMP-1 is an absolute requirement for collagen breakdown. Production of chimeric proteins combining features of MMP-1 and -3 as well as HPX domain deletion mutants showed complete abolishment of collagenolytic activity without any demonstrable effect on general hydrolytic activity towards gelatin, casein and other simple synthetic peptides (Murphy, Allan et al. 1992, Hirose, Patterson et al. 1993, Sanchez-Lopez, Alexander et al. 1993). Potentially complex substrate specificity determinants are controlled, in part, by the hemopexin domain. Certainly this is true for fibrillar collagen recognition, which requires complete full-length domain structure to ensure correct interplay between both the HPX and CAT domain.

## 1.6 Matrix Metalloproteinase Regulation

MMPs possess the ability to catabolise various matrix constituents necessary to alter cell fates and control developmental pathways. To prevent aberrant activity and subsequent overwhelming destruction within the ECM, exquisite regulatory mechanisms are in place (Figure 1.9).

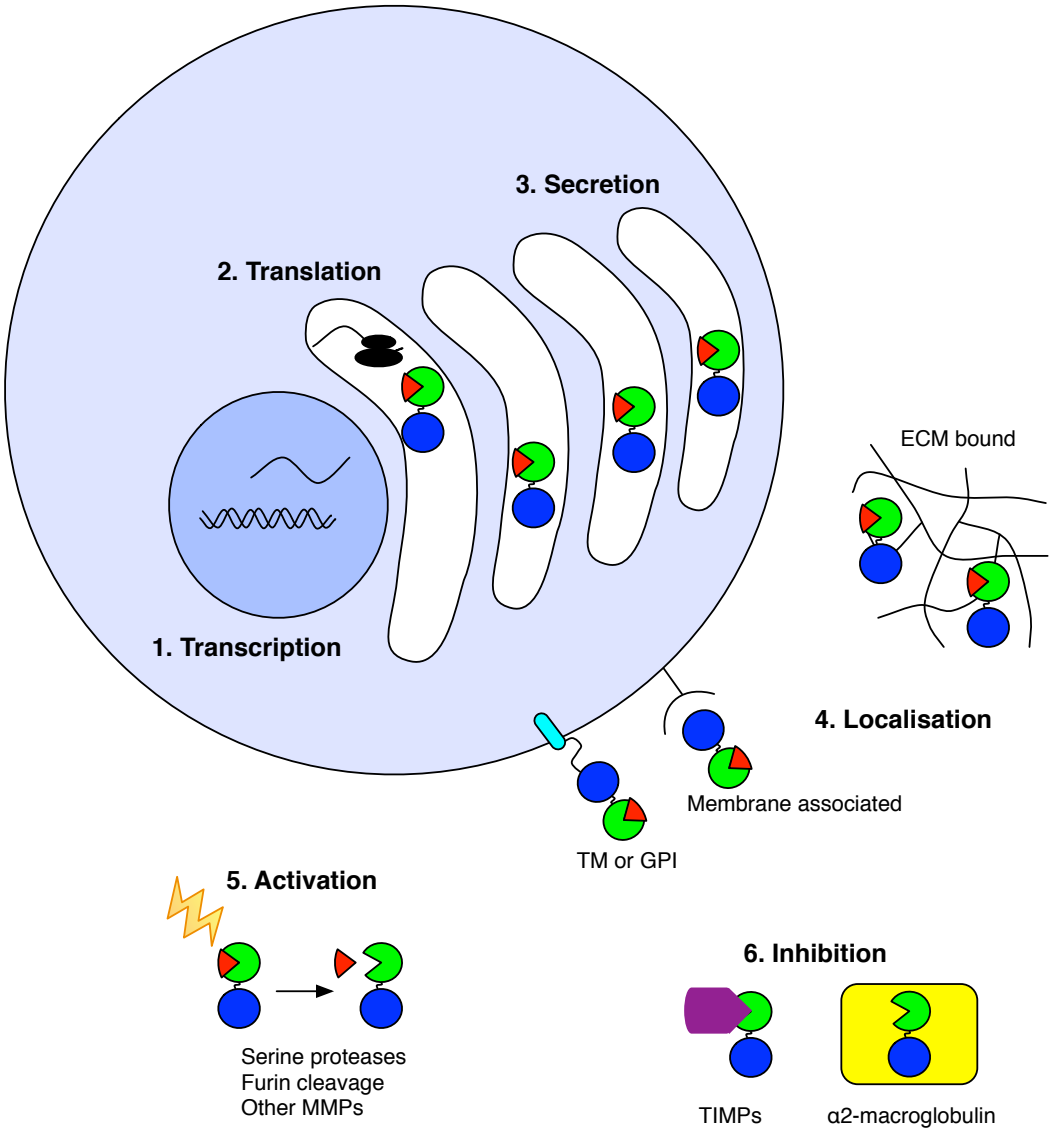



**Figure 1.9: Overview of MMP regulation.** Regulation occurs initially at both the (**1**) RNA transcription and (**2**) protein synthesis level. MMP activity is further controlled during (**3**) secretion through the *trans*-golgi network to the ECM, (**4**) localisation as seen with MMP-8 and membrane-bound MMPs, (**5**) zymogen activation and (**6**) inhibition by reversible endogenous inhibitors like TIMPs, or α2-macroglobulin which binds MMPs and leads to their clearance by endocytosis (Adapted from Page-McCaw, Ewald et al. 2007).

### 1.6.1 Transcriptional Regulation

Eight known human MMP genes are clustered on chromosome 11 at 11q21-23 with others distributed on chromosomes 1, 8, 12, 14, 16, 18, 20, 22. Extensive gene duplications in early vertebrate evolution particularly in the tetrapod lineage is responsible for the great protein structural divergence of MMPs (Huxley-Jones, Clarke et al. 2007). However, a number of MMP promoters share common features and as a result, genes are often transcribed simultaneously, in response to various inductive stimuli. Several *cis*-elements (e.g. activator protein (AP-1), polyoma enhancer A3 (PEA3), SP-1, beta-catenin/Tcf-4 and nuclear factor-kB (NF-kB) sites), and their corresponding trans-activators (Figure 1.10), have been implicated in regulating expression of MMP genes (Reviewed in Yan and Boyd 2006).

MMP promoters are grouped into three categories based on promoter composition. The first group of inducible promoters (MMP-1, -3, -7, -9, -10, -12, -13, -19 and -26) contain a TATA box and often a PEA3-binding site adjacent to an AP-1 binding site approximately ~70 bp upstream of the transcriptional start site. The second group (MMP-8, -11 and -21), lack the AP-1 site while the third group (MMP-2, -14 and -28) also lack the TATA box with constitutive expression mainly determined through binding of the GC box by the ubiquitous SP-1 family of transcription factors (Reviewed in Yan and Boyd 2006).

Basal MMP gene expression in normal resting tissue is usually low or even undetectable. Transient increases in transcription occur during normal matrix remodelling in response to signal transduction events that target and induce specific elements within the promoter region. For example, presence of a Tcf-4 site within the MMP-2, 7, -9 and -26 promoters mediates MMP expression in response to Wnt signalling (Crawford, Fingleton et al. 1999, Marchenko, Marchenko et al. 2004, Wu, Crampton et al. 2007), while possession of a NF-KB site within the MMP-9 promoter directs expression via an inflammatory cytokine-activated pathway (Bond, Fabunmi et al. 1998). The addition of MAPK, SMAD and STAT signalling as well as differential promoter response adds further levels of complexity and thereby directs combinatorial control of MMP promoters to generate diverse tissue-specific ECM environments.

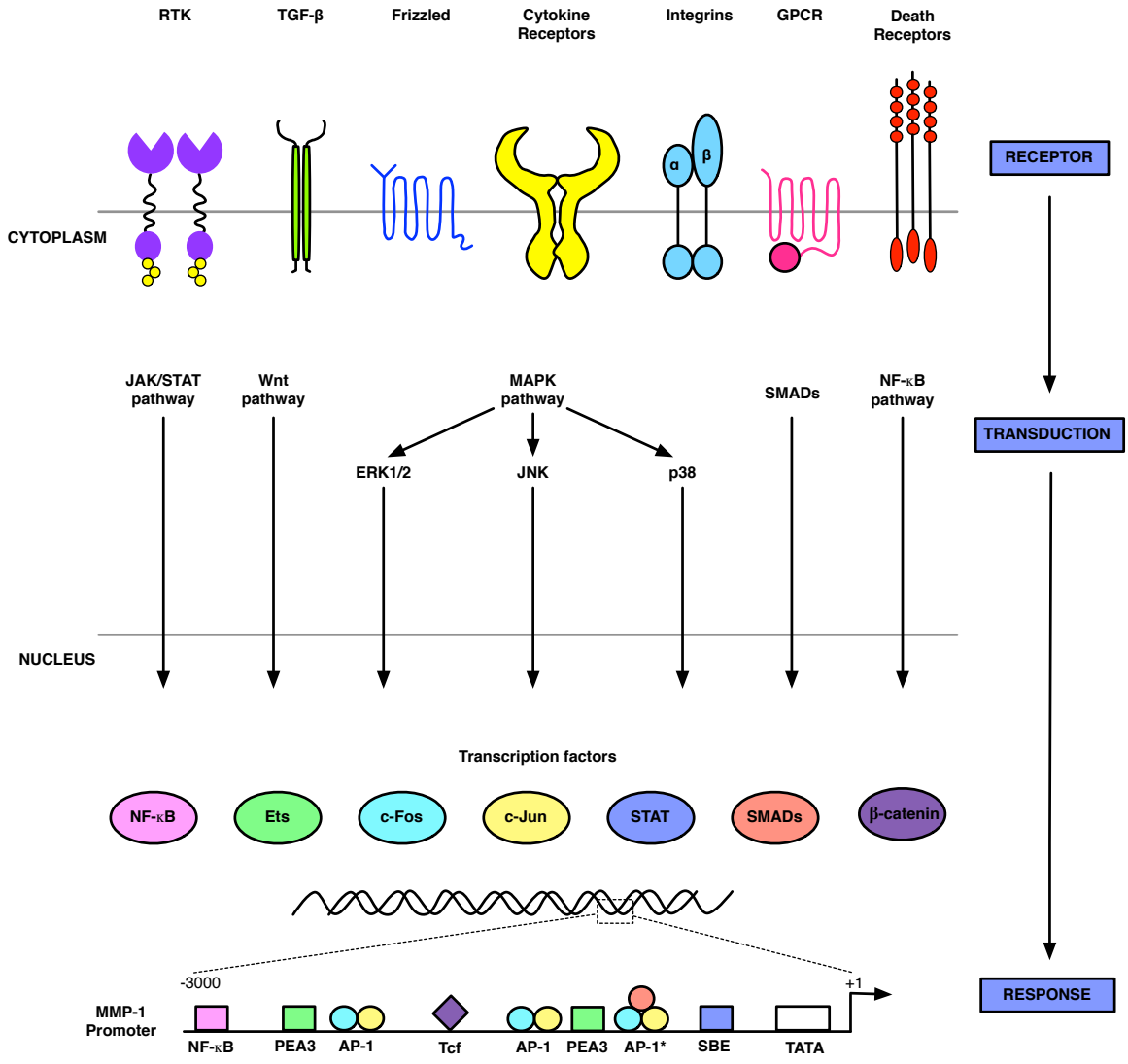



**Figure 1.10: Signal transduction pathways known to induce MMPs.** Ligand-receptor interactions occuring at the stromal-epithelial interface activate several signal transduction pathways. Spatial and temporal regulation can be achieved through integration of *trans*-activating factors with different combinations of *cis*-binding elements within the promoter region, in this case MMP-1. Relative positions of *cis*-elements are not drawn to scale. Abbreviations: RTK, receptor tyrosine kinase; TGF-β, transforming growth factor beta; GPCR, G-protein coupled receptor; JAK, Janus kinases; STAT, signal transducer and activator of transcription; MAPK, mitogen-activated protein kinase; ERK, extracellular signal-related kinase; JNK, c-Jun N-terminal activated kinases; NF-κB, nuclear-factor-κB; Ets, erythroblastosis twenty six; PEA3, polyomavirus enhancer-A binding protein site; AP-1, Activator protein-1 site; Tcf, T-cell factor site; SBE, STAT-binding element. (Adapted from Overall and Lopez-Otin 2002, Rowan and Young 2007).

### 1.6.2 Localisation

Direct tethering to the cell-surface via transmembrane and GPI-anchors is an obvious method of localisation achieved by the MT-MMP family, whereby membrane integration and tethering can ensure the focused proteolysis necessary for cellular motility and invasion of neighbouring tissues. For example, MMP-14 has been shown to

accumulate at the invadopodia of breast carcinoma cell line MDA-MB-231 (Artym, Zhang et al. 2006) as a result of membrane integration, which is evident by transmembrane and cytoplasmic domain truncations of MMP-14 that prevented invadopodial localisation (Nakahara, Howard et al. 1997, Lehti, Valtanen et al. 2000). MMP-14 has also been reported in lamellopodia of migrating osteoclasts and endothelial cells (Sato, del Carmen Ovejero et al. 1997, Galvez, Matias-Roman et al. 2001), this cellular migration appears to be mediated by adhesion protein CD44 (Mori, Tomari et al. 2002). Furthermore, CD44 receptor clustering is reported to facilitate cell-surface docking of MMP-7 and -9 (Bourguignon, Gunja-Smith et al. 1998, Yu and Stamenkovic 1999, Yu, Woessner et al. 2002, Desai, Rogers et al. 2007).

Other cell adhesion molecules heavily implicated in MMP localisation are the heterodimeric transmembrane integrins. Immunofluorescence and immunoprecipitation studies show co-localisation of MMP-14 with integrins αvβ3 and β1 (Ellerbroek, Fishman et al. 1999, Hofmann, Westphal et al. 2000, Deryugina, Ratnikov et al. 2001). It has been suggested that these integrins may have distinctly different roles, with β1 integrins acting as a reservoir for MMP-14 at endothelial cell junctions prior to redistribution to the αvβ3 integrins at the migrating cell front (Galvez, Matias-Roman et al. 2002). Similarly MMP-2 is known to bind the αvβ3 integrin (Brooks, Stromblad et al. 1996). This is mediated by the C-terminal HPX domain and is linked to vascular invasion during angiogenesis (Brooks, Stromblad et al. 1996, Brooks, Silletti et al. 1998, Silletti, Kessler et al. 2001). However other roles of this interaction may exist and include impairment of integrin signalling (Menon, Singh et al. 2006), presumably by occluding access of extracellular cognate signalling molecules; and potentially as an additional cell-surface activation mechanism (Hofmann, Westphal et al. 2000). MMP-1 also utilises its HPX domain as a contact point for specific α2β1 integrin-binding (Dumin, Dickeson et al. 2001, Stricker, Dumin et al. 2001). These particular integrins show enhanced pericellular expression at the leading edge of migrating human keratinocytes (Decline and Rousselle 2001), thereby spatially restricting MMP-1 proteolysis. Migration across type I collagen in the dermal matrix is dependent on the action of MMP-1 as demonstrated by a series of experiments which inhibited cellular motility by directly preventing collagenolytic activity, this included using α2 antibodies to block the integrin-MMP complex (Pilcher, Dumin et al. 1997).

### 1.6.3 Zymogen Activation

All secreted MMPs are thought to be released into the ECM in an inhibited zymogen form with enzyme latency maintained by the small globular PRO domain that occludes the active site cleft thus preventing substrate hydrolysis. Thiol modification, allosteric perturbation or stepwise processing of protease susceptible loop regions of the PRO domain are postulated to destabilise the domain. This is believed to disrupt the cysteine coordination of the zinc ion, thus triggering the 'cysteine switch' (Springman, Angleton et al. 1990, Van Wart and Birkedal-Hansen 1990) and exposing the autolytic cleavage site (Figure 1.11).

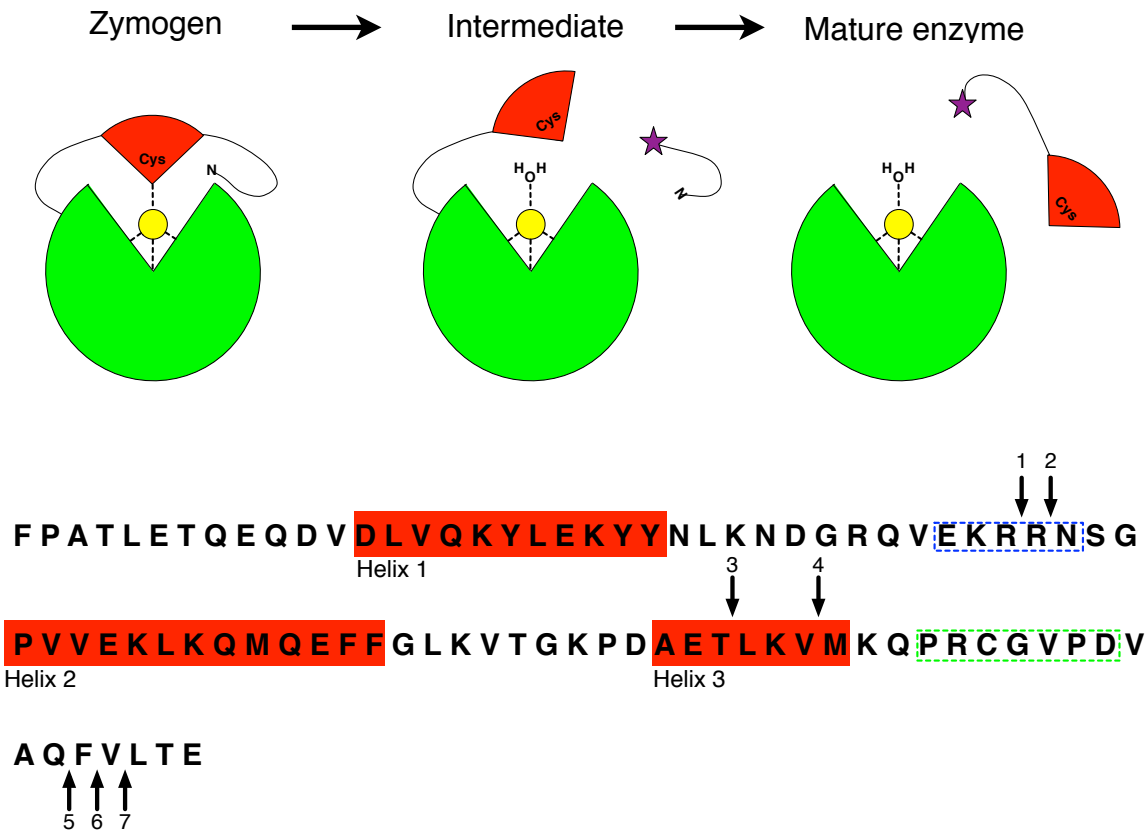



**Figure 1.11: Protease cleavage of the PRO domain.** Schematic illustration of enzymatic activity suppression by the PRO maturation by proteolytic propeptide removal, indicated by purple stars, which disrupt $Zn^{2+}$ chelation by the thiol moiety in the 'cysteine switch' PRCGVPD consensus sequence. Primary amino acid sequence of the MMP-1 PRO domain. Structural helix elements are shown in red. The bait region in loop one is indicated by the blue dashed box. The cysteine switch region is indicated by the green dashed box. The position of protease cleavage sites are indicated by arrows (**1**) trypsin site, (**2**) trypsin and plasma kallikrein (PKK) site, (**3**) MMP-1 autocatalytic site following treatment with PKK and plasmin, (**4**) MMP-1 autocatalytic site following treatment with APMA, (**5**) MMP-3 site, (**6**) MMP-1 autocatalytic site, (**7**) MMP-1 autocatalytic site following treatment with APMA. (Adapted from Birkedal-Hansen, Moore et al. 1993).

However, the complex physiological and biochemical activation mechanisms during enzyme maturation continue to remain largely elusive. Nevertheless, they are theorized to occur either at the cell surface via complexation with MT-MMPs mediated by endogenous MMP inhibitors, members of the TIMP family (Itoh, Binner et al. 1995, Overall, King et al. 1999, Overall, Tam et al. 2000, Morgunova, Tuuttila et al. 2002), or via an activation cascade initiated by serine proteases (Nagase, Enghild et al. 1990, Suzuki, Enghild et al. 1990).

Activation mechanisms are not limited to proteases. *In vitro* organomercurial compounds like 4-aminophenyl mercuric acetate (APMA), chaotropic ions, disulfide compounds, sulfhydryl alkylating agents and oxidants have all been shown to activate MMPs most likely through exposure, permanent modification or liberation of the inhibitory zinc-binding cysteine residue (Stricklin, Jeffrey et al. 1983, Nagase, Enghild et al. 1990, Rajagopalan, Meng et al. 1996, Gu, Kaul et al. 2002).

#### 1.6.3.1 Matrix Metalloproteinases

*In vitro*, many MMPs have been shown to participate in processing other proMMPs (Figure 1.12), by triggering the stepwise activation mechanism leading to autolysis and PRO domain removal.

Maturation of proMMP-2 has been well-studied and is achieved by TIMP-dependent activation. TIMP-2 binds membrane bound MT1-MMP using it's inhibitory N-terminal domain (Fernandez-catalan, Bode et al. 1998). Now immobilized, the TIMP-2 C-terminal domain binds the hemopexin domain of proMMP-2 through blades III and IV (Itoh, Binner et al. 1995, Overall, King et al. 1999, Overall, Tam et al. 2000). MT1-MMP homodimerization occurs (Itoh, Takamura et al. 2001) forming a tetrameric complex thereby allowing cleavage of the proMMP-2 scissile peptide bond between Asn37 and Leu38 (Will, Atkinson et al. 1996). A mature MMP-2 molecule with full gelatinase activity is produced following subsequent transactivation by MMP-2 (Atkinson, Crabbe et al. 1995). MT1-MMP has also been shown to activate proMMP-13 (Knauper, Bailey et al. 2002), although independently of TIMP-2.

MMP-3 has also been shown to be a potent activator of several MMPs including MMP-9 (Ogata, Enghild et al. 1992). For the collagenases, MMP-1 and MMP-8, N-terminal processing by MMP-3 results in enzyme 'superactivation' with a 4-12 fold increase in activity when compared to activation initiated by mercurials and other proteases (Murphy, Cockett et al. 1987, He, Wilhelm et al. 1989, Suzuki, Enghild et al.

1990, Knauper, Wilhelm et al. 1993). Thus, MMP-3 may play a key role in ECM degradation by initiating a cascade of MMP activity with subsequent degradation of the basement membrane and interstitial matrix via the collagenases MMP-1, -8 and -13, and MMP-9. However, the relevance of these superactivated enzyme forms and the physiological activation mechanism of most proMMPs remains unclear, particularly as a study using MMP-3 deficient mice revealed no impaired proMMP-9 activation (Lijnen, Silence et al. 1998).

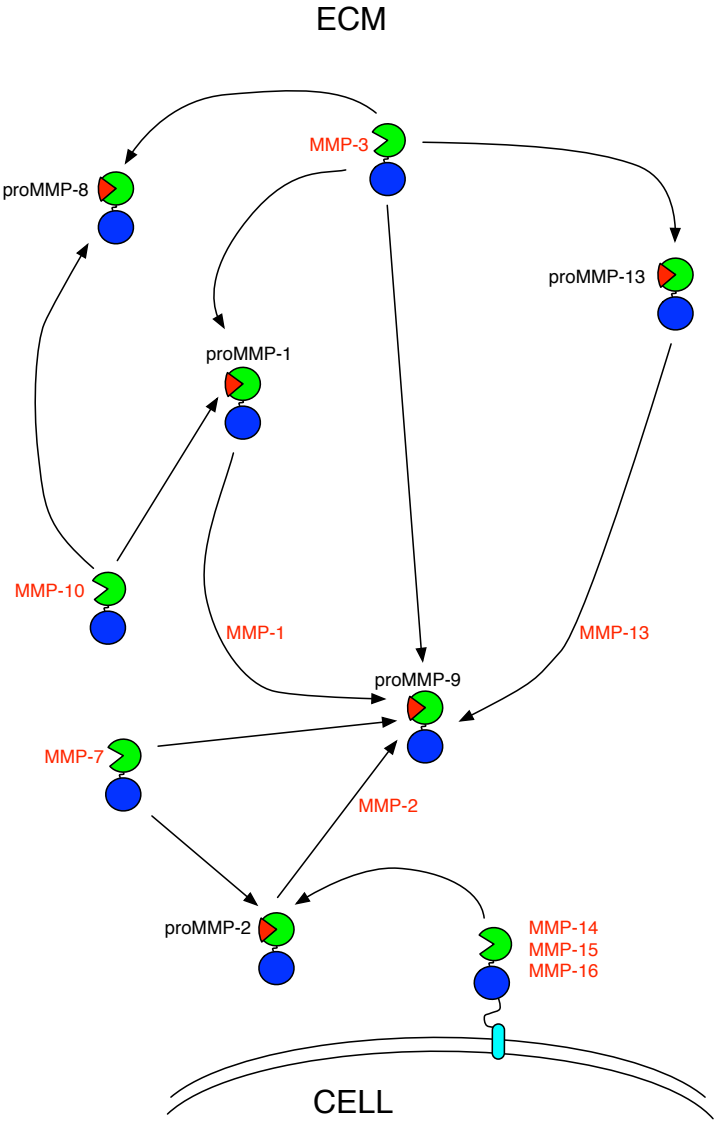



**Figure 1.12: MMP activation network.** Zymogen forms are labeled in *black* and active MMPs are labeled in *red*. The interactions depicted are not exhaustive. Data was compiled from (Ito and Nagase 1988, Nicholson, Murphy et al. 1989, Quantin, Murphy et al. 1989, Suzuki, Enghild et al. 1990, Nagase, Suzuki et al. 1992, Ogata, Enghild et al. 1992, Knauper, Wilhelm et al. 1993, Crabbe, O'Connell et al. 1994, Fridman, Toth et al. 1995, Sang, Birkedal-Hansen et al. 1995, Takino, Sato et al. 1995, Knauper, Murphy et al. 1996, Nakamura, Fujii et al. 1998, Ramos-DeSimone, Hahn-Dantona et al. 1999, Uria and Lopez-Otin 2000, Morrison, Butler et al. 2001).

### 1.6.3.2 Serine Proteases

To date, 178 human serine protease members have currently been identified (Quesada, Ordonez et al. 2009) and the vast majority of these enzymes can be grouped within four classic serine protease families S1, S8, S10, S14; typified by chymotrypsin, subtilisin, carboxypeptidase Y and Clp protease respectively (Reviewed in Hedstrom 2002). Known MMP activators urokinase-type-plasminogen activator (uPA), tissue-type-plasminogen activator (tPA), plasmin, plasma kallikrein and furin are all denoted as S1 serine proteases, i.e each possesses the characteristic Asp-His-Ser triad with preference for cleavage of substrates, dependent on arginine or lysine at position P1 (Rawlings and Barrett 1999). As previously mentioned, the general mechanism of MMP activation involving stepwise proteolytic processing begins within protease susceptible loop regions of the PRO domain. Hydrolysis occurs within basic amino acid motifs located within the 'bait' region of the solvent exposed loop between helix 1 and 2 (Nagase, Suzuki et al. 1991). More detail of the molecular basis of proteolytic activation has been revealed during a stopped-flow, freeze-quench, x-ray spectroscopic study of proMMP-9 activation catalysed by the serine protease tissue kallikrein. The temporal sequence of events leading to catalysis is apparently triggered during protein-protein complexation with activation occurring via a pentacoordination intermediate (contributions from Cys73, His218, His222, His228 and Glu219) transient $Zn^{2+}$ ion charge changes and proton transfer from the active site glutamate residue (Rosenblum, Meroueh et al. 2007). Intriguingly, this study concluded that cysteine dissociation occurs prior to significant proteolysis. MMP activity in the absence of PRO domain removal has previously been reported (Bannikov, Karelina et al. 2002), but whether this is a general feature of physiological relevance remains to be determined. However, what is clear, is that for many MMPs proteolytic processing forms unstable intermediates that are rapidly processed to the mature, active enzyme (Stricklin, Jeffrey et al. 1983, Okada, Harris et al. 1988, Murphy, Ward et al. 1989, Nagase, Enghild et al. 1990, Suzuki, Enghild et al. 1990, Suzuki, Lees et al. 1995, Knauper, Lopez-Otin et al. 1996, Knauper, Will et al. 1996, Monea, Lehti et al. 2002, Golubkov, Chekanov et al. 2007).

#### 1.6.3.2.1 uPA-plasmin System

The urokinase-type Plasminogen Activator (uPA)-plasmin system plays a pivotal role in matrix degradation by regulating the extracellular proteolytic cascade resulting in

plasmin formation. Proteolytic processing of ubiquitous zymogen plasminogen (Plg) by uPA or tPA results in activated plasmin, which has a distinct role in thrombolysis as well as being implicated as a potential physiological activator of proMMPs (Nagase, Enghild et al. 1990, Ramos-DeSimone, Hahn-Dantona et al. 1999, Monea, Lehti et al. 2002). Blood and vasculature restricted expression of tPA suggests that the uPA-plasmin system is more relevant for pericellular MMP activation, particularly as urokinase-type plasminogen activator receptor (uPAR) possesses a GPI-anchor ensuring the Plg activation cascade is localised to the cell surface.

The central binding cavity of uPAR is formed by a concave arrangement of three domains (Llinas, Le Du et al. 2005), which mediate high affinity binding (0.1-1 nM) with 1:1 stoichiometry to both zymogen (pro-uPA) and mature uPA (Cubellis, Nolli et al. 1986, Roldan, Cubellis et al. 1990, Moller, Pollanen et al. 1993). Specific binding to uPAR increases the cell surface concentration of active uPA, which in turn promotes uPAR clustering in the lipid-enriched microdomains of the plasma membrane (Cunningham, Andolfo et al. 2003). Plg, in contrast, is restricted to the cell surface by a large group of heterogeneous receptors, which in combination contribute to the high cellular Plg binding capacity estimated at $10^4$-$10^7$ binding sites per cell (Hajjar, Harpel et al. 1986, Ganz, Dupuis et al. 1991). Reported candidate Plg-receptors include actin, cytokeratin (Hembrough, Li et al. 1996), α-enolase (Redlitz, Fowler et al. 1995) and integrin αVβ3, which typically bind in a lysine-dependent manner (Ellis, Whawell et al. 1999).

As well as providing protection from circulating inhibitors (Plow, Freaney et al. 1986, Redlitz, Fowler et al. 1995) α2-antiplasmin and the serpins, plasminogen activator inhibitor 1 (PAI-1) and PA-2, binding to cell-surface receptors appears to promote a conformationally altered activation-susceptible Plg form (Christensen and Molgaard 1991, Andronicos and Ranson 2001). Presumably close spatial proximity between the uPA-uPAR complex and receptor-bound Plg then stimulates cleavage of the Arg561-Val562 Plg peptide bond (Robbins, Summaria et al. 1967) resulting in enzymatic maturation and subsequent conversion of surface-bound plasminogen to plasmin. Plg activation is accelerated in the presence of receptor-bound uPA, while another study observed colocalisation between integral membrane receptor Plg-$R_{KT}$ and uPAR (Andronicos, Chen et al. 2010).

Cell surface-associated plasmin is known to process *in vitro* a number of latent MMPs including MMP-1, -2, -3, -9, -13 and -14 (Grant, Eisen et al. 1987, He, Wilhelm et al. 1989, DeClerck, Yean et al. 1991, Knauper, Will et al. 1996, Mazzieri, Masiero et

al. 1997, Okumura, Sato et al. 1997, Ramos-DeSimone, Hahn-Dantona et al. 1999, Santala, Saarinen et al. 1999, Monea, Lehti et al. 2002, Liu, Li et al. 2005). In all cases the PRO domain is targeted and processed by plasmin leading to decreases in molecular mass. Murine knockout experiments studying the effects of Plg-deficiency on wound-healing identified phenotypic overlap or synergy between the two enzyme systems (Lund, Romer et al. 1999, Lund, Green et al. 2006, Juncker-Jensen and Lund 2011) with decreased keratinocyte migration during reepithelialisation (Romer, Bugge et al. 1996, Green, Almholt et al. 2008). A similar study showed defective placental vascularization, intrauterine growth retardation and increased embryonic lethality in $Plg^{-/-}$ mice treated with broad spectrum MMP inhibitor galardin (Solberg, Rinkenberger et al. 2003). In a mouse model of inflammatory skin disease Bullous pemphigold, MMP-9 activation was found to be plasmin-dependent during the early stages of sub-epidermal blistering (Liu, Li et al. 2005). However, further support for the role of plasmin as a physiological activator in the MMP activation pathway has been lacking.

#### 1.6.3.2.2 Furin

Furin is a ubiquitously expressed member of the proprotein convertase (PC) family, a group of mammalian calcium-dependent subtilisin/kexin-like serine proteases primarily distributed in the *trans*-golgi network (Bosshart, Humphrey et al. 1994, Molloy, Anderson et al. 1999) that cleave proteins traversing constitutive secretory routes. The furin-like PCs; PC1, PC2, PACE4, PC4, furin, PC5A, PC5B and PC7 (Reviewed in Thomas 2002), are structurally related to yeast *Saccharomyces cerevisiae* endoproteinase Kex2 (Nakayama 1997) and recognise dibasic consensus sequences RXRR/RRKR/RRRR in protein precursors (Molloy, Bresnahan et al. 1992, Walker, Molloy et al. 1994, Rawlings and Barrett 1999). The stringent requirements of furin for arginine at P1 and P4 positions and the preference for polybasic substrate segments is the result of surface loops extending into the active site and negative electrostatic potential of subsite surfaces (Henrich, Cameron et al. 2003).

A number of MMPs, including MT-MMPs, are activated upon cleavage of the furin consensus sequence, which is located at the C-terminal end of the PRO domain. Furin-deficient cell lines, LoVo and RPE.40, have been repeatedly used to show the predominant role of furin as an activator of certain MMPs, of these MMP-14 is certainly the best characterised. Two potential proprotein convertase tetrabasic cleavage motifs reside within the MMP-14 PRO domain, R89-R-P-R↓C93 and R108-R-K-

R↓Y112 (Yana and Weiss 2000), however the first site exists as part of the cysteine switch motif (Remacle, Rozanov et al. 2006) and therefore enzyme maturation is the result of cleavage generating the Y112 N-terminus (Strongin, Collier et al. 1995, Lehti, Lohi et al. 1998, Osenkowski, Toth et al. 2004). Furthermore, Golubkov and colleagues have shown that maturation cleavage is facilitated by intradomain proteolytic processing at the P47-G-D↓L50 site (Golubkov, Chekanov et al. 2007). Intriguingly, in the absence of this processing, the MMP-14 PRO domain excised by furin retained potent inhibitory activity with a $k_i$ of 160-200 nM (Golubkov, Cieplak et al. 2010). Retention of a metalloproteinase inhibitory domain has been observed previously. Following furin processing in the trans-Golgi network the PRO domain of ADAM12 remained non-covalently associated with the mature enzyme (Wewer, Morgelin et al. 2006). Therefore, it is unlikely that furin alone plays a significant role in PRO domain destabilisation and MMP maturation.

#### 1.6.3.3 Other Mechanisms

Organomercurial compounds are routinely employed to activate latent MMPs *in vitro*, and are thought to initiate the stepwise autolytic excision of the PRO domain. It has been proposed that compounds like APMA directly react with the $Zn^{2+}$-chelating cysteine residue leading to displacement of the sulfhydryl group. However, CD data suggests that APMA stimulates enzyme maturation through PRO domain perturbation rather than direct cysteine modification, at least for MMP-3 (Chen, Noelken et al. 1993). As yet, the mode of action by non-physiological agents, like APMA, remains unknown.

In contrast, oxidative mechanisms are the likely instigator of MMP activation *in vivo,* along with serine proteases. Increases in MMP expression and activation in hepatic stellate cells, human monocytes and human coronary smooth muscle cells has been observed during periods of oxidative stress (Galli, Svegliati-Baroni et al. 2005, Lu and Wahl 2005, Valentin, Bueb et al. 2005). Furthermore, the physiological nitrous oxide (NO) donor S-nitrosocysteine has been shown to activate proMMP-9 by S-nitrosylation of the cysteine during cerebral ischemia (Gu, Kaul et al. 2002).

Reactive oxygen and nitrogen species are known to readily react with thiol moieties. Fu and colleagues previously demonstrated proMMP-7 activation after subsequent treatment with the cytotoxic oxidant hypochlorous acid (HOCl), a derivative of hydrogen peroxide ($H_2O_2$). In this study, liquid chromatography electrospray

ionization mass spectrometry (LC/MS) was used to show thiol residue conversion to sulfinic acid, which preceeded increased proteolytic activity (Fu, Kassim et al. 2001). HOCl is produced following the release of myeloperoxidase (MPO) from the cytoplasmic granules of neutrophils and monocytes during phagocytosis (Bergt, Marsche et al. 2001). Consequently, agents known to inhibit the MPO-$H_2O_2$ system or scavenge the resulting chlorinated oxidants have been shown to significantly reduce MMP activity (Peppin and Weiss 1986, Meli, Christen et al. 2003) and are therefore of interest when considering MMPs and their role in inflammatory response.

### 1.6.4 Endogenous Inhibitors

A variety of molecules have been implicated in MMP inhibition. The abundant non-specific endopeptidase inhibitor, plasma α2 macroglobulin (α2M), is known to bind irreversibly to all active MMPs. Upon cleavage of 'bait domains' within α2M, a structural change results in covalent endopeptidase capture along with exposure of receptor domains, which facilitate rapid scavenger receptor-mediated endocytosis and clearance (Enghild, Salvesen et al. 1989, Sottrup-Jensen and Birkedal-Hansen 1989, Sottrup-Jensen and Birkedal-Hansen 1992, Moestrup, Holtet et al. 1993). Thrombospondin-1 binds progelatinases preventing enzymatic activation and RECK, a GPI-anchored glycoprotein, has been shown to downregulate levels of MMP-9 (Takahashi, Sheng et al. 1998, Oh, Takahashi et al. 2001). However, the best characterised and principal endogenous inhibitors are members of the broad spectrum TIMP family.

#### 1.6.4.1 TIMPs

Four TIMPS have been identified in vertebrates, and together they can preferentially suppress the MMP degenerative potential in connective tissue through tight reversible binding with MMPs in a 1:1 stoichiometry (Cawston, Murphy et al. 1983, Welgus, Jeffrey et al. 1985, Overall 1994).

TIMP-1, -2, -3 and -4 have six disulphide bonds forming the inhibitory N-terminal and C-terminal domains (Williamson, Marston et al. 1990). The complex crystal structures of N-TIMP-1 with the catalytic domains of MMP-1 and MMP-3 (Figure 1.13) show the coordination of the catalytic zinc via the α-amino nitrogen and the peptide carbonyl group of the strictly conserved TIMP cysteine residue at position 1.

Inhibition is maintained through interaction with the glutamate residue leading to expulsion of the catalytic water molecule. Further interactions, with the active site cleft and the $S_1$' specificity pocket, bind and block the active site (Gomis-Ruth, Maskos et al. 1997, Iyer, Wei et al. 2007).

Understanding the structural basis of TIMP and MMP interactions is essential for future development of therapeutic inhibitors for diseases with unbalanced degradation of ECM components as a consequence of MMP activity. Currently nearly all synthetic inhibitors incorporate hydroxamic acid, carboxylate or thiol group to chelate the catalytic zinc ion.

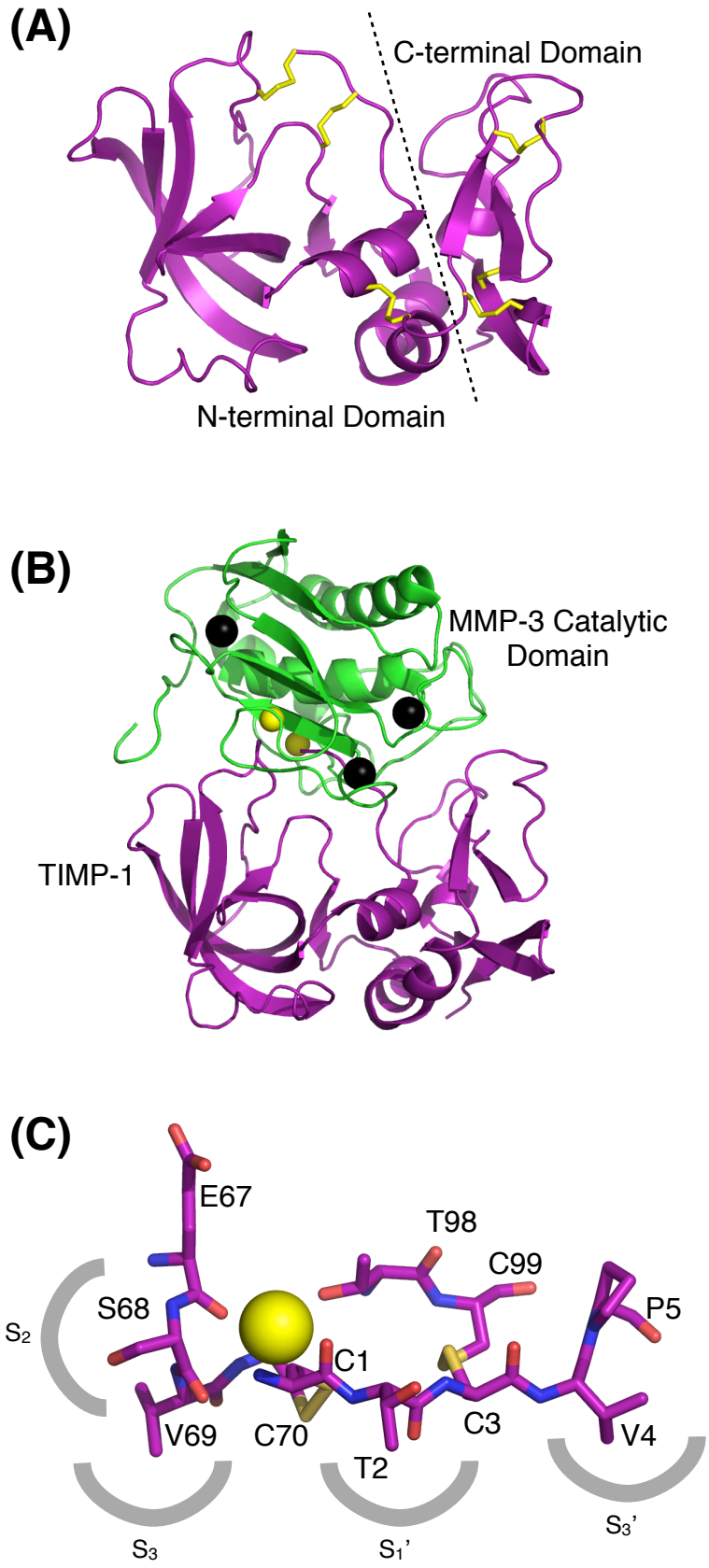



**Figure 1.13 : Structure of TIMP-1 and its interaction with MMP-3.** (**A**) Ribbon representation of the crystal structure of TIMP-1, showing locations of disulphide bonds (*yellow sticks*). (**B**) TIMP-1 (*purple*) bound to the MMP-3 catalytic domain (*green*), zinc ions are depicted as *yellow* spheres and calcium ions are coloured *black* (PDB accession code - 1UEA). (**C**) The primary TIMP-1 interaction site, subsites within the MMP active site cleft are indicated (*grey*). (Adapted from Brew and Nagase 2010).

## 1.7 The Paradox of Collagenolysis

How collagenases recognise and hydrolyse collagen is not clearly understood. Despite the presence of 31 potential cleavage sites in the triple helical regions of type I, II, III and IV collagens (Fields 1991), interstitial collagenases process the collagen triple helix with precision at a single locus generating characteristic ¼ and ¾ length fragments. Analysis suggests this specificity is conferred by imino-rich sequences preceeding the scissile bond followed by a loose helical region within the subsequent four triplets, as well as the presence of an arginine residue at position P'5 or P'8 (Fields 1991). However, proteolytic cleavage of collagen by collagenases is inhibited by localisation of this unique cleavage site within a narrow cleft, created both by microfibril formation and the presence of the C-telopeptide (Perumal, Antipova et al. 2008). Of the total number of procollagen monomers present in a collagen fibril (50 nm diameter), only 10-12 % are surface exposed and accessibility is further restricted due to the polymerization of these fibrils, which may account for the modest cleavage rate of 22 collagen molecules degraded per collagenase per hour (Welgus, Jeffrey et al. 1980).

Further still is the conflicting size dimension of enzyme and substrate, which typically prevents peptide bond hydrolysis. Tropocollagen (15 Å diameter) is substantially larger than the active site presented by MMP-1 (5 Å), which is able to accommodate individual polypeptide chains only (Chung, Yoshida et al. 2004). Studies using type III collagen-like peptides also reveal that collagen scissile bonds are not solvent exposed and are positioned at least 7 Å from the catalytic $Zn^{2+}$ ion, and therefore remain inaccessible to collagenases (Stultz 2002, Chung, Yoshida et al. 2004).

It has been postulated that either the active site undergoes large structural rearrangements to accommodate the triple helical substrate, or collagen needs to be unwound for a single polypeptide chain to be delivered to the active site. Unwinding may occur spontaneously or be induced by collagenases.

Arguably, unfolded vulnerable collagen states exist. NMR studies and molecular dynamics (MD) simulations suggest an absence of the triple helical fold near the collagenase cleavage site (Fiori, Sacca et al. 2002, Nerenberg and Stultz 2008). However, these vulnerable regions are generally resistant to degradation when treated with isolated collagenase CAT domains (Schnierer, Kleine et al. 1993, Chung, Yoshida et al. 2004, Han, Makareeva et al. 2010). A recent study explored type I collagen degradation by MMP-1 and -8 CAT domains, demonstrating the redundancy of the HPX domain (Salsas-Escat, Nerenberg et al. 2010). While degradation was observed, it

was significantly less than that observed for full-length MMP-1, requiring high enzyme concentrations and long incubation times. Therefore, the HPX domain, which has been proposed to either unwind or stabilise conformationally-labile substrate states, is essential for efficient collagenolysis.

## 1.8 Collagen-binding by the HPX domain

It has long been suspected that non-catalytic domains are involved in collagenolysis through either the conformational selection mechanism or via substrate processing and preparation prior to cleavage. Both the HPX domain of collagenases and the FN2 domains of gelatinases have been implicated in this process and are thought to possess exosites with essential roles in collagen binding.

The presence of the HPX domain has been shown as an absolute requirement for catalysis and has been proposed to facilitate bond presentation through substrate positioning via these exosites (Lauer-Fields, Juska et al. 2002).

Topology of the HPX domain is generally equivalent for the vast majority of MMPs with the final C-terminal ~200 residues constructing the typical four blade β-propeller arrangement (Li, Brick et al. 1995, Gomis-Ruth, Gohlke et al. 1996, Jozic, Bourenkov et al. 2005, Iyer, Visse et al. 2006, Bertini, Calderone et al. 2008, Tochowicz, Goettig et al. 2010). In MMP-1 the domain is stabilised by the presence of an intradomain disulphide bond between blades 1 (Cys259) and 4 (Cys447) (Jozic, Bourenkov et al. 2005, Iyer, Visse et al. 2006). Substitution of these cysteine residues disrupts the HPX domain of MMP-1 sufficiently, resulting in complete loss of type I collagenolysis (Windsor, Birkedal-Hansen et al. 1991). Total HPX domain deletion mutants have equivalent reduction in collagenolytic ability but with no effect on general catalytic competence with hydrolytic activity toward gelatin, casein and simple synthetic peptides maintained (Clark and Cawston 1989, Murphy, Allan et al. 1992).

Circular dichroism spectroscopical analysis of triple helical collagen in the presence of the HPX domain suggests non-catalytic domain (NCD) induced perturbation of collagen substrate (Tam, Moore et al. 2004, Gioia, Monaco et al. 2007) and triple helicase activity mediated by the HPX domain of MMP-1 has been demonstrated with local unwinding of the collagen chains through preferential binding of α2(I) chain of Type I collagen (Chung, Yoshida et al. 2004). This appears to be necessary for scissile bond exposure and sequential polypeptide chain cleavage. The molecular basis of this interaction remains to be elucidated, however exosite regions

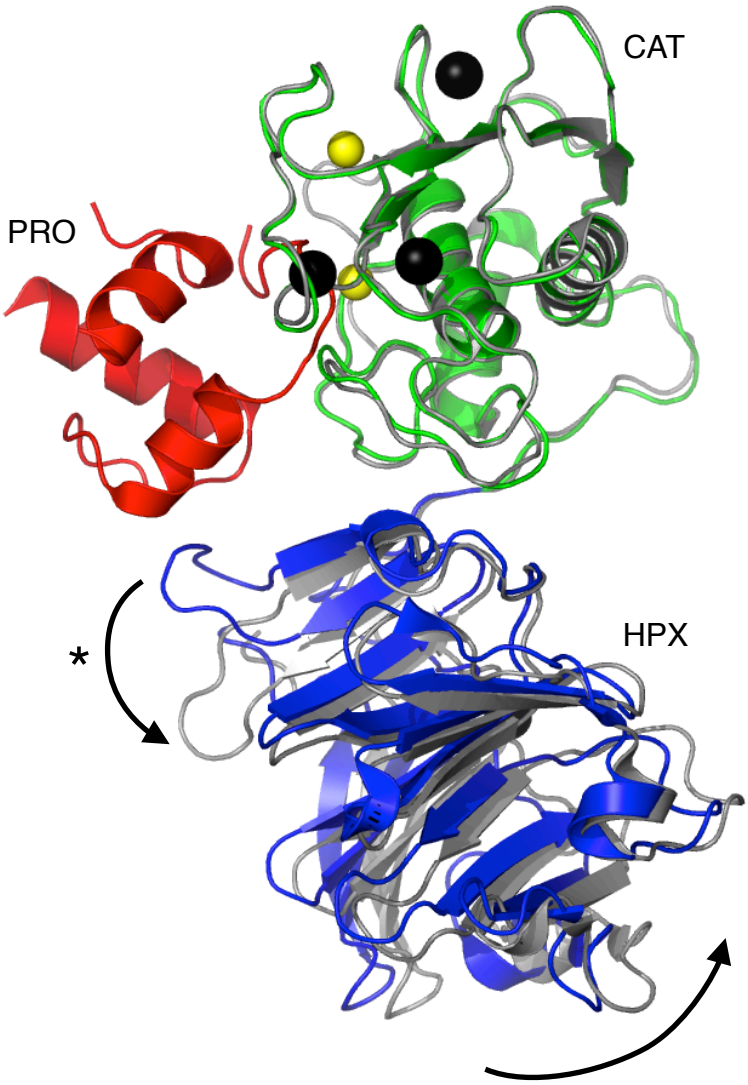



**Figure 1.14: Conformational differences in the HPX domain of proMMP-1 and MMP-1.** PRO-1, CAT-1 and HPX-1 domains are shown in *red*, *green* and *blue* respectively. Zymogen and mature MMP-1 have been superimposed using their catalytic domains (RMSD 0.815). The curved arrows indicates the relative movement upon activation. The CD-loop is denoted by an asterisk (*). Coordinates taken from PDB depositions (1SU3 and 2CLT). Image rendered in Pymol Molecular Graphics System (version 1.5.0.4.), Schrödinger, LLC.

Table 1.3: HPX interactions with PRO domain residues

| MMP-1 residues | PRO domain residues | | Shift in Cα position (Å) |
|---|---|---|---|
| | van der Waals interactions | Hydrogen bonds | |
| Phe308 | Lys74, Asp80, Glu82 | - | 15.52 |
| Tyr309 | Gly72, Leu73, Glu82 | Tyr309 OH- Glu82 OE1 (2.6 Å) | 11.88 |
| Pro310 | Gly72, Leu73, Lys74 | - | 8.87 |

(Adapted from Iyer, Visse et al. 2006)

involved in collagen recognition and binding have been identified by hydrogen/deuterium exchange mass spectrometry (HD/MS) experiments. Subsequent mutagenesis of residues Ile290 and Arg291, result in decreased affinity for collagen (Lauer-Fields, Chalmers et al. 2009). However, HPX domain deletion studies of interstitial collagenase MMP-13 have shown that catabolic activity towards collagens IV, IX, X and XIV, fibronectin and tenascin is maintained indicative of exosite absence in some MMPs.

Domain replacement studies are revealing. Collagenolytic activity is abolished in catalytically-competent chimeric proteins comprised of either the MMP-1 or -8 CAT domain fused with the HPX domain from members of the stromelysin family (Murphy, Allan et al. 1992, Hirose, Patterson et al. 1993, Sanchez-Lopez, Alexander et al. 1993). Despite type I collagen binding activity by the MMP-3 HPX domain (Murphy, Allan et al. 1992), fusion of this domain is insufficient to restore collagenase activity. Therefore, collaboration between correctly paired collagenase domains is necessary for collagenolysis. This is supported in a study by Chung and colleagues, which found high concentrations of isolated MMP-1 CAT and HPX domains was required for collagen breakdown, but connection of the two domain components by the linker was necessary for a more effective enzyme (Chung, Yoshida et al. 2004).

## 1.9 MMP Domain Rearrangements

Non-catalytic domains clearly play a major role in fibrillar collagen binding however little is known about the collaborative mechanism between domains which facilitates collagen catabolism. Significant conformational changes subsequent to propeptide cleavage, allowing enzyme maturation, have been observed. Superimposition of the human proMMP-1 structure (PDB accession code: 1SU3) with MMP-1 (PDB accession code: 2CLT) reveals significant changes in conformation. Removal of the PRO domain appears to disrupt a number of hydrophobic interactions unmasking residues Phe308-Pro310 located within the loop region between β-strand C and D on blade I of the hemopexin domain. The CD-loop undergoes a notable rearrangement of position (Figure 1.14) with the largest shift observed in the Phe308 Cα position (Table 1.3). It has been suggested that HPX displacement is mediated by Arg300, which acts as a pivot (Jozic, Bourenkov et al. 2005). However, changes in domain orientation may also result

from conformational freedom of the variable-length linker region, which ranges from 14 to the 68 amino acid residues observed for MMP-9.

Collagenase linkers are known to play a role in collagen degradation. De Souza et al. (1996) suggested a role for the flexible linker region as a potential intercalating agent driving the unwinding process of collagen through insertion between the polypeptide chains supported by interactions with the HPX domain. However, a chimeric protein composed of the N-terminal MMP-3 domain combined with the C-terminal region of MMP-1, including the linker, was not endowed with collagenolytic activity (Murphy, Allan et al. 1992). Instead, extensive mutagenesis and assay of MMP-1 linker residues identified the importance of a glycine residue at position 272 (P270-I-**G**-P-Q-T275). Substitution with asparagine at this position was found to reduce type I collagen breakdown to 13% without affecting substrate specificity and binding, and consequently it was proposed that the linker is more likely to govern spatial domain arrangements through a hinge-bending motion (Tsukada and Pourmotabbed 2002). Furthermore, such conformational freedom may be a general feature of MMPs. Solution studies of MMP-9 and -12 support the existence of multiple enzyme conformations, which transiently assume more elongated conformations (Rosenblum, Cohen et al. 2007, Bertini, Calderone et al. 2008). More recently, interdomain flexibility has been observed for MMP-1, although less pronounced than that of MMP-9 and -12 (Bertini, Fragai et al. 2009). It remains unclear if collagenase flexibility is a requirement for triple helical substrate recognition, manipulation and breakdown.

## 1.10 Aims and Objectives

The success of collagenolysis relies on determinants in both the triple helical collagen substrate and the collagenase enzyme. MMP-1 induces collagenolysis and previous studies have implicated it's HPX domain in binding and destabilising the collagen substrate in preparation for hydrolysis by the CAT domain. Furthermore, studies of full-length MMPs support a hypothesis in which a structurally dynamic enzyme is able to utilise a collaborative domain mechanism to recognise and degrade collagen. Therefore, the primary purpose of this study is to both explore HPX domain mediated collagen recognition and the postulated flexible state of MMP-1, with an aim to facilitate understanding of the collagenolytic mechanism. Initially this involved the expression of recombinant human MMP proteins from clones provided by Dr Rob Visse and Prof. Hideaki Nagase (Imperial College, London) in *E. coli,* and the optimised

purification of refolded proteins to produce yields sufficient for structural and functional characterisation. This work is described in detail in Chapter 2, along with structural assessment of isolated domains by nuclear magnetic resonance (NMR) and limited proteolysis, and initial characterisation of hydrolytic activity using substrate cleavage assays.

The precise location of HPX exosite(s) has, until recently, remained largely elusive. However, work conducted in the laboratory of Dr Andrew Pickford (University of Portsmouth) has identified a putative collagen-binding site within the HPX domain of MMP-1. An extensive alanine-scanning site-directed mutagenesis strategy is combined with Surface Plasmon Resonance (SPR) to characterise this site. This work is described in Chapter 3.

Conformational freedom and transient separation of the CAT and HPX domain of MMPs has been observed during NMR and SAXS studies, consequently interdomain flexibility has been proposed to play a role in the collagenolytic mechanism. Therefore, the role of the dislocated and compact state of MMP-1 will be elucidated using SAXS undertaken at the Diamond Light Source (Harwell Science and Innovation Campus, Didcot) and SPR, the results of which are described in Chapters 4 and 5.

Comparison of the zymogen and mature MMP-1 crystal structures highlights the potential role of PRO-HPX contacts, which may restrain domain mobility and/or prohibit substrate recognition and proteolysis by maintaining enzyme latency. In Chapter 6 mutagenesis is used to disrupt the PRO-HPX interface, SAXS is used to investigate the global zymogen form and SPR is used to assess collagen recognition by these mutants.

During the latter stages of this project, two studies were published that provide significant insight into collagen catabolism by collagenases. The first study describes a mechanism of collagenolysis that is derived from NMR experimental data of MMPs and MMP-peptide complexes, and docking experiments using protein-protein docking programme HADDOCK (Bertini, Fragai et al. 2012). The second study used biochemical experiments and the first crystallographic structure of MMP-1 in complex with a triple helical peptide to show that collagenolysis relies on multiple exosite interactions (Manka, Carafoli et al. 2012). These studies are discussed in detail within Chapter 7, and the results contained within this thesis are considered in the context of these latest findings.

# CHAPTER 2: MMP PRODUCTION

## 2.1 Introduction

Protein expression strategies exploit prokaryotic or eukaryotic cellular machinery to synthesise heterologous functional proteins for proteomic study. The expression systems available are highly variable (Table 2.1) and selection is typically dependent on the physical properties and function of the target protein. However, laboratory resources, production cost and downstream experimental applications may often impact host system selection.

Table 2.1: Comparison of Expression Systems

| Desired characteristics | Bacteria | Yeast | Insect | Mammalian cell culture |
|---|---|---|---|---|
| Cell growth | Rapid | Rapid | Slow | Slow |
| Complexity of growth medium | Minimum | Minimum | Complex | Complex |
| Cost of growth medium | Low | Low | High | High |
| Expression level | High | Low-High | Low-High | Low-Moderate |
| Extracellular expression | Secretion to periplasm | Secretion to medium | Secretion to medium | Secretion to medium |
| Protein folding | Refolding usually required | Refolding may be required | Proper folding | Proper folding |
| N-linked glycosylation | None | High mannose | Simple, no sialic acid | Complex |
| O-linked glycosylation | No | Yes | Yes | Yes |
| Phosphorylation | No | Yes | Yes | Yes |
| Acetylation | No | Yes | Yes | Yes |
| Acylation | No | Yes | Yes | Yes |
| γ-Carboxylation | No | No | No | Yes |

Reproduced from Gene Expression Systems: Using Nature for the Art of Expression.
Edited by: Joseph M. Fernandez and James P. Hoeffler © 1999 Academic Press., San Diego, CA

Recombinant MMP proteins have been successfully expressed in chinese hamster ovary cells (Housley, Baumann et al. 1993), in the methylotrophic yeast *Pichia pastoris* (Rosenfeld, Ross et al. 1996) and in the Sf9 insect cell line using the baculovirus expression system (del Mar Barbacid, Fernandez-Resa et al. 1998). However, *Escherichia coli* remains the most widely used prokaryotic expression host for many recombinant proteins and MMP production is no exception (Young, Rowan et al. 2010). *E. coli* is often the preferred system for large-scale heterologous protein production. In fact, 80% of all solved protein structures submitted to the PDB in 2003, and nine of the thirty one approved therapeutic proteins in the subsequent three years, were prepared in *E.coli* (Sorensen and Mortensen 2005, Walsh 2006). Rapid, high-density cultivation on inexpensive substrates, high heterologous protein yields, well-established isotopic labelling protocols and versatile expression vector systems ensure its continued popularity.

The pET System (Novagen) includes over 42 commercially available plasmid expression vector types and 15 accompanying bacterial host strains developed for efficient target protein production (MerckMillipore n.d.). Derived from cloning vector pBR322, the pET plasmids (Studier and Moffatt 1986) are almost exclusively for use in *E. coli* λDE3 lysogen strains, that contain a phage construct encoding T7 RNA polymerase (T7RNAP) under the control of the lacUV5 promoter (Figure 2.1). Target genes are transcriptionally silent until T7RNAP expression is induced by the lactose analogue, isopropyl-1-thio-β-D-galactopyranoside (IPTG). T7RNAP is highly efficient and promoter-specific (Chamberlin and Ring 1973, Cheetham, Jeruzalmi et al. 1999). Consequently, the translational machinery of *E. coli* is rapidly saturated by the transcribed mRNA, generating target protein levels exceeding 50% of total cellular protein content within hours (Studier and Moffatt 1986). However, heterologous protein expression may be severely diminished as a result of codon bias, where synonymous codons are used with varying species-dependent frequencies (Grantham, Gautier et al. 1980, Sharp, Cowe et al. 1988). Rare *E. coli* codons include AGG, AGA and CGA encoding arginine, and CTA, ATA and CCC encoding leucine, isoleucine and proline respectively (Data obtained through the website http://www.kazusa.or.jp/codon/ Nakamura, Gojobori et al. 2000). Deficiencies in cognate tRNAs, which are more commonly found in eukaryotic genomes, may lead to enhanced mRNA degradation, amino acid misincorporation, frame shifts and stalled translation (Reviewed in Kane 1995). Bacterial codon bias-adjusted BL21 derivatives have been devised to maximise expression. Commercially available host strains include Rosetta (DE3) cells and Agilent

Technologies' BL21-CodonPlus (DE3) series which contain extra copies of the *E. coli* argU, ileY, and leuW tRNA genes to overcome codon bias of the heterologous target gene.

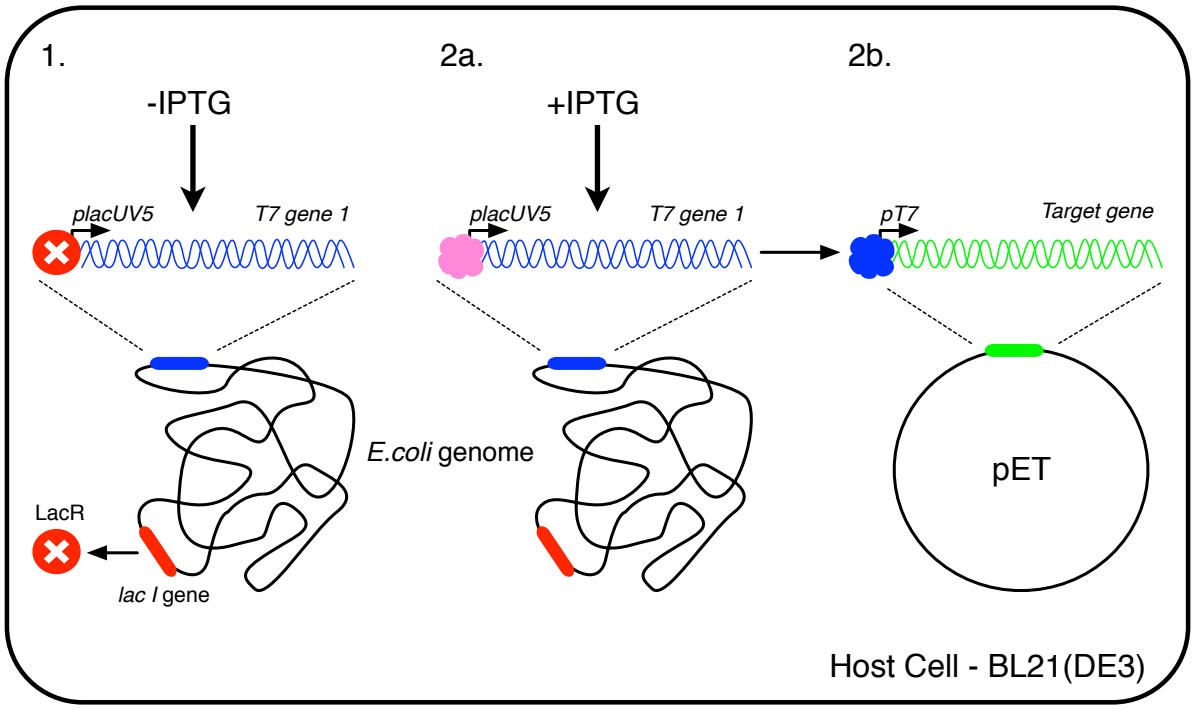



**Figure 2.1: Vector and host elements of the prokaryotic pET expression system**. An *E. coli* strain containing a chromosomal copy of the T7 RNA polymerase (*T7 gene 1*) is transformed with a recombinant plasmid containing the gene of interest. (**1**) *T7 gene 1* transcription is repressed by the *lac* repressor (LacR) bound to its cognate operator within the *lacUV5* promoter. (**2a**) Transcription of *T7 gene 1* and subsequent expression of T7 RNA polymerase (T7RNAP) is stimulated using IPTG, an analogue of allolactose, which displaces LacR from the operator allowing binding by *E. coli* RNA polymerase (*pink*). (**2b**) Induced production of T7RNAP (*blue*) transcribes the target gene (*green*) by binding specifically to the T7 viral promoter (*pT7*). (Adapted from Csun.edu n.d.)

Further complications ensue as protein expression progresses. *E. coli* has a limited capacity for complex posttranslational modifications, and inefficiently forms disulphide bonds. Therefore, enhanced heterologous protein yields often result in insoluble inclusion bodies (Bowden, Paredes et al. 1991) formed almost exclusively from aggregated target protein (Rinas and Bailey 1992). Nevertheless, recovery of bioactive proteins is possible and inclusion body (IB) deposition provides protection from host proteases, and a convenient method for purification of over-expressed proteins. A high specific density of ~1.3 g/ml$^{-1}$ (Margreiter, Messner et al. 2008) means IBs can be isolated from the cell homogenate by centrifugation using moderate rotor speeds, while

development of efficient solubilisation procedures using denaturants such as urea and guanidine hydrochloride, can produce IB preparations of high purity. Acquisition of protein functionality relies on suitable refolding methods, which reduce the propensity for aggregation via hydrophobic intermolecular interactions while promoting formation of the native state. Common strategies include dilution or dialysis to remove detergents and denaturants.

### 2.1.1 Nuclear Magnetic Resonance

Production of high-quality heterologous proteins amenable for further structural and biochemical studies is essential. Determining sample quality is often achieved using protein Nuclear Magnetic Resonance (NMR) spectroscopy to assess the structural nature of the target protein. While a detailed description of NMR spectroscopy is beyond the scope of this thesis, a brief description of the NMR phenomenon is provided.

Atomic nuclei possess the intrinsic quantum property of angular momentum, termed 'spin', and a magnetic moment. Nuclei with an odd mass number (e.g. $^{1}H$, $^{13}C$, $^{15}N$, $^{19}F$) have fractional spins. Nuclei with a ½ integer spin quantum value (I), enables the existence of two energy spin states when an external magnetic field is applied; a low energy α ground state ($I = +½$) and a high energy β state ($I = -½$). An unequal population of the two possible quantum energy states in thermal equilibrium results in bulk sample magnetization in the direction (*z*-axis) of the applied field. The spins precess at a resonance frequency determined by the individual nuclear gyromagnetic ratio (Table 2.2) multiplied by the strength of the applied magnetic field. Application of radio-frequency electromagnetic radiation at this frequency stimulates transitions between the energy states rotating the magnetization vector onto the transverse plane. Magnetization in this plane decays exponentially as the spin system returns to thermal equilibrium and is detected as a Free Induction Decay (FID). A fourier transformation is then applied to the FID to obtain the NMR spectrum.

The precise resonant frequency of the energy transition is directly proportional to the strength of the applied magnetic field. However, variations in local magnetic 'shielding' by valence electrons, due to the surrounding chemical environment, affect the precessional frequency of nuclei and, thus, the radiation required to stimulate the energy transition. Differences in this resonance frequency is termed 'chemical shift' (δ) and can be measured relative to a standard compound.

Table 2.2: Properties of selected spin -½ nuclei

| Isotope | Natural abundance[a] (%) | Magnetic moment (μ/μN) | Gyromagnetic ratio ($\gamma/10^7$ rad $s^{-1}$ $T^{-1}$) |
|---|---|---|---|
| $^{1}H$ | 99.9885 | 4.837353570 | 26.7522128 |
| $^{13}C$ | 1.07 | 1.216613 | 6.728284 |
| $^{15}N$ | 0.368 | -0.49049746 | -2.71261804 |
| $^{19}F$ | 100 | 4.553333 | 25.18148 |
| $^{31}P$ | 100 | 1.95999 | 10.8394 |

[a] representative isotopic compositions
(Taken from Harris, Becker et al. 2001)

Local deviations in shielding resulting from the proximity of adjacent atoms in folded proteins can easily be determined from the dispersed NMR spectra of two dimensional heteronuclear chemical shift correlation spectra. The $^{1}H$, $^{15}N$ Heteronuclear Single Quantum Correlation (HSQC) experiment used in this study, exploits the strong scalar spin-spin coupling ($^{1}J^{HN}$ = ~92 Hz) between directly bonded amide proton and nitrogen nuclei of $^{15}N$ isotopically-enriched proteins, to assess the native structure of MMP-1 domains.

### 2.1.2 Site-Directed Mutagenesis

Alanine-scanning mutagenesis is a powerful technique for determining the catalytic or functional role of individual residues in a protein. Alanine is the most popular residue for substitution since it does not alter the main-chain conformation, unlike glycine and proline, and possesses a non-reactive aliphatic sidechain, a single methyl group. Therefore, substitution with alanine allows sidechain deletion effects of protein-ligand complex affinity to be measured while attempting to preserve native tertiary structure.

Using *in vitro* site-directed mutagenesis (SDM), oligonucleotide primers containing the desired mutation, in this case the alanine codon, can be incorporated into the amplicon thereby changing the original nucleotide sequence and consequently the translated amino acid sequence. The QuikChange® XL Site-Directed Mutagenesis Kit

(Agilent Technologies) provides a rapid and efficient means of manipulating and modifying *E. coli* plasmid DNA containing the gene of interest (Figure 2.2).

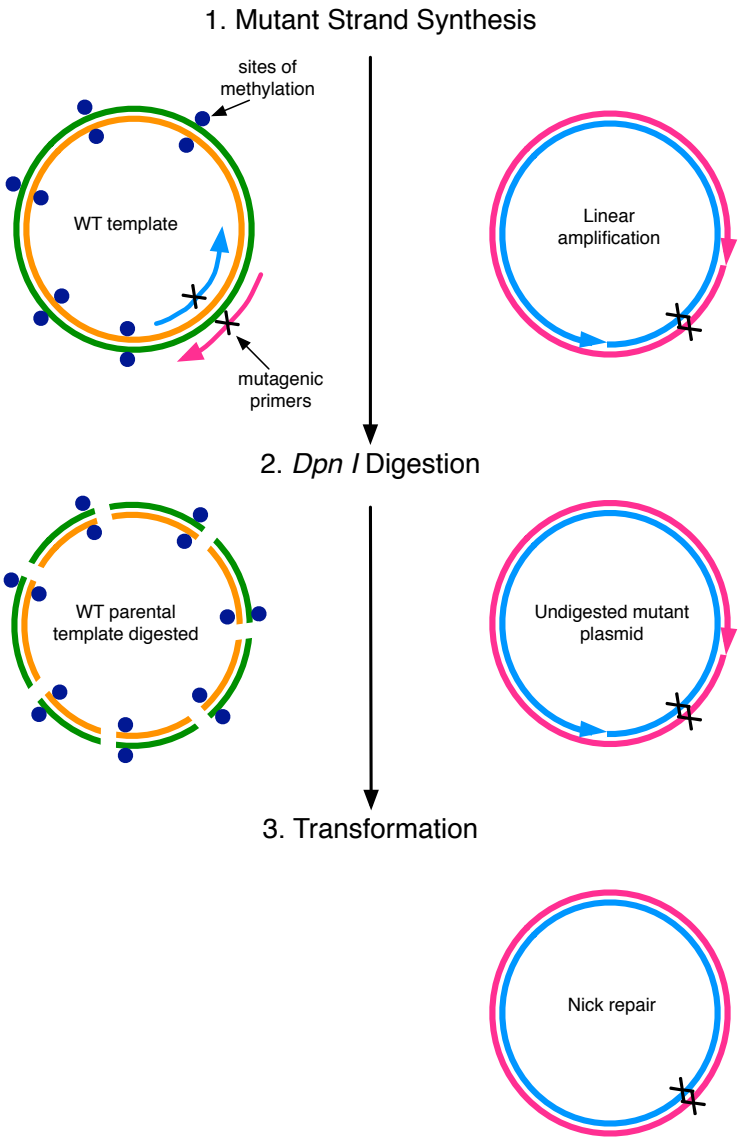



**Figure 2.2: *In vitro* site-directed mutagenesis using the QuikChange® XL Site-Directed Mutagenesis Kit (Agilent Technologies)**. Thermal cycling is utilised to denature the DNA template, anneal mutagenic primers and extend the primers using *Pfu*Ultra HF DNA polymerase during Mutant Strand Synthesis (1). The original template is digested with *Dpn I*, which targets GATC methylated and hemimethylated sites, shown as *dark blue circles* (2). The remaining mutated plasmid is transformed into competent cells for nick repair (3). (Adapted from Agilent Technologies n.d.)

In brief, this method uses a thermostable high fidelity DNA polymerase, *Pfu*Ultra HF, for mutagenic oligonucleotide primer-directed plasmid replication. *Pfu*Ultra HF

lacks terminal transferase activity but possesses 3'-5' exonuclease activity, as a consequence this enzyme synthesises DNA with significantly higher fidelity than *Taq* DNA polymerase. Mutant strand synthesis is achieved by annealing complementary primers incorporating the desired mutation during temperature cycling in the presence of *Pfu*Ultra HF, which generates a mutated plasmid with staggered nicks. Treatment with *Dpn I,* an endonuclease specific for a methylated restriction site 5´-GATC-3´, which is a target in *dam*+ strains of *E. coli,* is used to selectively digest the parental DNA template. The isolated *in vitro* synthesised 'mutant' plasmid is then transformed into competent cells for nick repair.

Mutagenic efficiency is determined using pWhitescript (4.5 kb) as a control plasmid, which contains a premature stop codon (TAA) in place of glutamine codon (CAA) within the β-galactosidase gene of pBluescript II SK(-) phagemid. This disrupts β-galactosidase hydrolysis of 5'bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal), an analogue of lactose, producing white colonies. Mutagenic oligonucleotide control primers incorporate a point mutation that reverts the C residue in the glutamine codon, restoring β-galactosidase activity. Subsequent cleavage of X-gal yields 5'bromo-4-chloro-3-hydroxyindole and galactose, the former product dimerises and is oxidized into the blue product 5,5'-dibromo-4,4'-dichloro-indigo. Transformed XL10-Gold Ultracompetent cells can then be screened for β-galactosidase activity as demonstrated by a blue colony phenotype. However, non-control reactions require DNA sequencing to ensure the desired mutation has been successfully incorporated.

This chapter describes the expression and purification of the recombinant human proteins proMMP-3ΔHPX (hereafter referred to as proCAT-3), proMMP-1 (and mutants thereof), HPX-1 (and mutants thereof) and proMMP-1ΔHPX (hereafter referred to as proCAT-1*) using the pET vector system, and the subsequent analysis of protein quality using both NMR and substrate cleavage assays.

## 2.2 Materials

### 2.2.1 Chemical and Biochemical Reagents

Except where stated otherwise, all chemical and biochemical reagents were supplied by Fisher Scientific. $^{15}N$ ammonium sulphate for isotopic labeling was supplied by Cambridge Isotope Laboratories, Inc. Protogel 37.5:1 Acrylamide:Bisacrylamide stabilised solution (30% total monomer) for SDS-PAGE was supplied by National Diagnostics. SeeBlue Plus2 Pre-Stained Standard for protein molecular weight visualisation was supplied by Life Technologies. COmplete, EDTA-free Protease Inhibitor Cocktail Tablets for general protease inhibition was supplied by Roche Applied Science. IEX Sepharose™ Fast Flow chromatography media and columns were supplied by GE Healthcare Life Sciences. Urea (99.9%), cystamine dihydrochloride (97%) and deoxycholic acid (98.5%) were supplied by Acros Organics. Centrifugal filter units for protein concentration were supplied by Millipore. Standard Regenerated Cellulose Spectra/Por dialysis membrane was supplied by Spectrum Laboratories, Inc. StrataClean resins were supplied by Agilent Technologies. MMP substrates Ac-Pro-Leu-Gly-[2-mercapto-4-methyl-pentanoyl]-Leu-Gly-$OC_2H_5$ and Mca-Arg-Pro-Lys-Pro-Val-Glu-Nva-Trp-Arg-Lys-(Dnp)-$NH_2$ were purchased from Enzo Life Sciences and Bachem, respectively. Mca (7-methoxycoumarin-4-acetic acid) and Dnp (2,4-dinitroaniline) are a common donor/quencher pair incorporated into FRET-based fluorescent probes.

### 2.2.2 Nucleic Acids

The original pET-3a plasmid constructs containing the sequences for proMMP-3ΔHPX, proMMP-1*, HPX-1 and proMMP-1*ΔHPX cloned between NdeI and BamHI were generously donated by Dr Rob Visse and Prof. Hideaki Nagase, Imperial College London (See Appendix). The proMMP-1ΔHPX and proMMP-3ΔHPX inserts were subcloned into the equivalent sites of pET-28b by Dr Andy Pickford (University of Portsmouth), thus endowing the proteins with an N-terminal $His_6$ fusion tag.

### 2.2.3 Cell Strains

The *E. coli* DH5α strain [*E. coli* $F^-$ endA1 glnV44 thi-1 recA1 gyrA96 deoR nupG

Φ80d*lacZ*ΔM15 Δ(*lacZYA-argF*)U169, hsdR17($R_K^-$ $m_K^-$), *λ*-] was used for plasmid preparation.

XL10-Gold ultracompetent cells [*E.coli* TetrΔ (*mcrA*)*183* Δ(*mcrCB-hsdSMR-mrr*)*173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac* Hte [F' proAB *lacI*q*Z*Δ*M15* Tn*10* (Tetr) Amy Camr] were provided as part of the Quikchange II XL Site-Directed Mutagenesis Kit supplied by Agilent Technologies, for use in nick repair following mutagenesis and blue-white colour screening.

The *E. coli* strain BL21 (DE3) [*E. coli* B $F^-$ *dcm ompT hsdS*($r_B^-$ $m_B^-$) *gal λ*(DE3)] was a general protein expression strain used that lacks both the *lon* protease and the *ompT* outer membrane protease, which can degrade proteins during purification.

The *E. coli* strain BL21-CodonPlus® (DE3)-RIPL [*E. coli* B $F^-$ *ompT hsdS*($r_B^-$ $m_B^-$) $dcm^+$ $Tet^r$ *gal λ*(DE3) *endA* Hte [*argU proL* $Cam^r$] [*argU ileY leuW* Strep/$Spec^r$] was routinely used for protein expression of proMMP-1, HPX-1 and proMMP-3ΔHPX.

### 2.2.4 Buffer and Media Recipes

All buffers and growth media were made with 18.2 MΩ deionised water (Purite) and autoclaved at 120°C for 20 minutes at 15 psi, except for those containing heat-sensitive components, which were filter-sterilised using 0.22 μm filters (Millipore).

## 2.3 Methods

Due to the number of recombinant proteins expressed and purified during the course of this investigation (Table 2.3), the methods described herein remain non-specific.

Table 2.3: Recombinant proteins produced for study

| Protein | Name | Description |
|---|---|---|
| proMMP-3ΔHPX | proCAT-3 | C-terminally truncated MMP-3 zymogen form |
| MMP-3ΔHPX | CAT-3 | C-terminally truncated catalytically active MMP-3 |
| HPX-1 | HPX-1 | Isolated C-terminal hemopexin domain of wild-type MMP-1 |
| | HPX-1 R300A | Arginine → Alanine mutation at position 300 in the hemopexin domain |
| | HPX-1 F301A | Phenylalanine → Alanine mutation at position 301 in the hemopexin domain |
| | HPX-1 P307A | Proline → Alanine mutation at position 307 in the hemopexin domain |
| | HPX-1 Y309A | Tyrosine → Alanine mutation at position 309 in the hemopexin domain |
| | HPX-1 F316A | Phenylalanine → Alanine mutation at position 316 in the hemopexin domain |
| | HPX-1 S318A | Serine → Alanine mutation at position 318 in the hemopexin domain |
| | HPX-1 V319A | Valine → Alanine mutation at position 319 in the hemopexin domain |
| | HPX-1 P325A | Proline → Alanine mutation at position 325 in the hemopexin domain |
| | HPX-1 R337A | Arginine → Alanine mutation at position 337 in the hemopexin domain |
| | HPX-1 D338A | Aspartate → Alanine mutation at position 338 in the hemopexin domain |
| | HPX-1 Q354A | Glutamine → Alanine mutation at position 354 in the hemopexin domain |

Table 2.3: Recombinant proteins produced for study

| | | |
|---|---|---|
| | HPX-1 H358A | Histidine → Alanine mutation at position 358 in the hemopexin domain |
| | HPX-1 P361A | Proline → Alanine mutation at position 361 in the hemopexin domain |
| proMMP-1 | proMMP-1 | wild-type full-length zymogen |
| | proMMP-1 R300A | wild-type full-length zymogen with Arginine → Alanine mutation at position 300 in the hemopexin domain |
| | proMMP-1 F301A | wild-type full-length zymogen with Phenylalanine → Alanine mutation at position 301 in the hemopexin domain |
| | proMMP-1 F316A | wild-type full-length zymogen with Phenylalanine → Alanine mutation at position 316 in the hemopexin domain |
| | proMMP-1 S243C, S318C | wild-type full-length zymogen with Serine → Cysteine mutations at position 243 in the catalytic domain and 318 in the hemopexin domain |
| MMP-1 (active enzyme) | MMP-1 | wild-type full-length mature enzyme |
| | MMP-1 R300A | wild-type full-length mature enzyme with Arginine → Alanine mutation at position 300 in the hemopexin domain |
| | MMP-1 F301A | wild-type full-length mature enzyme with Phenylalanine → Alanine mutation at position 301 in the hemopexin domain |
| | MMP-1 F316A | wild-type full-length mature enzyme with Phenylalanine → Alanine mutation at position 316 in the hemopexin domain |
| | MMP-1 S243C, S318C | wild-type full-length mature enzyme with Serine → Cysteine mutations at position 243 in the catalytic domain and 318 in the hemopexin domain |
| proMMP-1ΔHPX (hydrolytically impaired) | proCAT-1* | Inactive C-terminally truncated MMP-1 zymogen form with Glutamine → Alanine mutation at position 200 in the active site of the catalytic domain |
| MMP-1ΔHPX (active enzyme) | CAT-1 | C-terminally truncated MMP-1 active enzyme form |

Table 2.3: Recombinant proteins produced for study

| | | |
|---|---|---|
| proMMP-1* (hydrolytically impaired) | proMMP-1* | Inactive full-length zymogen form with Glutamine → Alanine mutation at position 200 in the active site of the catalytic domain |
| | proMMP-1* R300A | Inactive full-length zymogen form with Glutamine → Alanine mutation at position 200 in the active site of the catalytic domain and Arginine → Alanine mutation at position 300 in the hemopexin domain |
| | proMMP-1* F301A | Inactive full-length zymogen form with Glutamine → Alanine mutation at position 200 in the active site of the catalytic domain and Phenylalanine → Alanine mutation at position 301 in the hemopexin domain |
| | proMMP-1* P307A | Inactive full-length zymogen form with Glutamine → Alanine mutation at position 200 in the active site of the catalytic domain and Proline → Alanine mutation at position 307 in the hemopexin domain |
| | proMMP-1* F308A | Inactive full-length zymogen form with Glutamine → Alanine mutation at position 200 in the active site of the catalytic domain and Phenylalanine → Alanine mutation at position 308 in the hemopexin domain |
| | proMMP-1* Y309A | Inactive full-length zymogen form with Glutamine → Alanine mutation at position 200 in the active site of the catalytic domain and Tyrosine → Alanine mutation at position 309 in the hemopexin domain |
| | proMMP-1* P310A | Inactive full-length zymogen form with Glutamine → Alanine mutation at position 200 in the active site of the catalytic domain and Proline → Alanine mutation at position 310 in the hemopexin domain |

Table 2.3: Recombinant proteins produced for study

| | | |
|---|---|---|
| | proMMP-1* F316A | Inactive full-length zymogen form with Glutamine → Alanine mutation at position 200 in the active site of the catalytic domain and Phenylalanine → Alanine mutation at position 316 in the hemopexin domain |
| | proMMP-1* S243C, S318C | Inactive full-length zymogen form with Glutamine → Alanine mutation at position 200 in the active site of the catalytic domain and Serine → Cysteine mutations at position 243 (catalytic domain) and 318 in the hemopexin domain |
| MMP-1* (hydrolytically impaired) | MMP-1* | Inactive mature form with Glutamine → Alanine mutation at position 200 in the active site of the catalytic domain |
| | MMP-1* R300A | Inactive mature form with Glutamine → Alanine mutation at position 200 in the active site of the catalytic domain and Arginine → Alanine mutation at position 300 in the hemopexin domain |
| | MMP-1* F301A | Inactive mature form with Glutamine → Alanine mutation at position 200 in the active site of the catalytic domain and Phenylalanine → Alanine mutation at position 301 in the hemopexin domain |
| | MMP-1* P307A | Inactive mature form with Glutamine → Alanine mutation at position 200 in the active site of the catalytic domain and Proline → Alanine mutation at position 307 in the hemopexin domain |
| | MMP-1* F308A | Inactive mature form with Glutamine → Alanine mutation at position 200 in the active site of the catalytic domain and Phenylalanine → Alanine mutation at position 308 in the hemopexin domain |
| | MMP-1* Y309A | Inactive mature form with Glutamine → Alanine mutation at position 200 in the active site of the catalytic domain and Tyrosine → Alanine mutation at position 309 in the hemopexin domain |

Table 2.3: Recombinant proteins produced for study

| | |
|---|---|
| MMP-1* P310A | Inactive mature form with Glutamine → Alanine mutation at position 200 in the active site of the catalytic domain and Proline → Alanine mutation at position 310 in the hemopexin domain |
| MMP-1* F316A | Inactive mature form with Glutamine → Alanine mutation at position 200 in the active site of the catalytic domain and Phenylalanine → Alanine mutation at position 316 in the hemopexin domain |
| proMMP-1* S243C, S318C | Inactive mature form with Glutamine → Alanine mutation at position 200 in the active site of the catalytic domain and Serine → Cysteine mutations at position 243 (catalytic domain) and 318 in the hemopexin domain |

Construct-specific methods are described in Tables 2.5-2.8 and contain details of both the purification pathway required and conditions necessary for optimal production of the recombinant protein of interest.

### 2.3.1 Preparation of *E. coli* strains for chemical transformation

*E.coli* bacterial strains frozen and stored at -80°C were streaked out on Lysogeny Broth (LB) agar plates and incubated at 37°C for 12-16 hours. Single, well-isolated colonies were selected and used to inoculate 10 mL of pre-mixed LB Miller liquid media (components per L: 10 g Tryptone, 5 g Yeast extract, 10 g NaCl, pH 7.0 at 25 °C) and incubated overnight at 37°C with shaking at 250 rpm using an Innova 4400 shaking incubator.

A baffled flask containing 250 mL of LB liquid media was inoculated with 250 µL of the overnight culture and cells were grown at 37°C with shaking at 250 rpm until an $OD_{600}$ of 0.6 was reached. Growth rate was monitored at 600 nm using a Bioware CO8000 Cell Density Meter. The culture was decanted into a pre-chilled 500 mL flat-bottomed centrifuge bottle and cells were harvested by centrifugation at 5000 rpm for 10 minutes at 4°C in a Beckman Coulter Allegra 25R refrigerated centrifuge (rotor

AT-14-10). The supernatant was discarded and the remaining pellet was gently resuspended in 80 mL of ice-cold CCMB80 buffer (10 mM KOAc, 80 mM $CaCl_2.2H_2O$, 20 mM $MnCl_2.4H_2O$, 10 mM $MgCl_2.6H_2O$, 10% (v/v) glycerol, adjusted to pH 6.4 with 1 mM HCl, filter-sterilised and stored at 4°C) and incubated on ice for 20 minutes. Cells were re-centrifuged and supernatant discarded as previously described, prior to resuspension in 10 mL of ice-cold CCMB80 buffer. 250 µL of resuspended cells were mixed with 1 mL of liquid LB media and cell density was assessed as previously described. Chilled CCMB80 was added to harvested cells to yield a final $OD_{600}$ of 1-1.5. These chemically competent cells were routinely stored in 250 µL aliquots at -80°C for 6-12 months without notable effects on transformation efficiency.

### 2.3.2 Transformation of chemically competent *E. coli* strains

An aliquot of chemically competent cells was thawed on ice and 50 µL of this cell suspension was decanted into a pre-chilled 15 mL polypropylene tube. 1 µL of the appropriate plasmid DNA at a concentration of 10 ng/µL was added to the cell suspension and incubated for 30 minutes on ice. The transformation reaction was heat-pulsed for 45 seconds in a 42°C water bath and immediately cooled on ice for an additional 2 minutes. 900 µL of liquid LB media was added to the transformation reaction prior to incubation at 37°C with shaking at 225 rpm for 1 hour to allow cell recovery and expression of antibiotic resistance. 200 µL of the reaction suspension was spread onto an LB agar plate supplemented with appropriate antibiotics for selection purposes (100 µg/mL ampicillin for pET-3a selection, 25 µg/mL kanamycin for pET-28b selection and 34 µg/mL chloramphenicol to maintain the pACYC plasmid in the BL21-CodonPlus® (DE3)-RIPL cell strain). Plates were inverted and incubated at 37°C for 12-16 hours. Typically this transformation method generated efficiencies of $>1x10^6$ cfu/µg of plasmid DNA.

### 2.3.3 Plasmid DNA purification and maintenance

DNA plasmid templates were extracted from *E. coli* DH5α cells using the QIAprep Spin Miniprep Kits for plasmid DNA purification supplied by QIAGEN, which uses the modified alkaline lysis method of Birnboim and Doly (1979).

A single transformed DH5α colony was used to inoculate 10 mL of LB liquid

media supplemented with appropriate antibiotic(s) (100 µg/mL ampicillin, 50 µg/mL kanamycin, 34 µg/mL chloramphenicol) and was incubated at 37°C for 12-16 hours with shaking (250rpm). Bacterial cells were harvested by centrifugation in 15 mL polypropylene tubes at 4000 rpm for 10 minutes at 4°C using a Beckman Coulter Allegra 25R refrigerated centrifuge (rotor TS-5.1-500) and the QIAprep miniprep procedure was followed (Qiagen 2006). In brief, cleared bacterial cell lysate was applied and reversibly bound to the QIAprep silica membrane in the presence of chaotrophic salts. After extensive washing, purified plasmid DNA was eluted in nuclease-free dd$H_2O$ (Purite)

Plasmid DNA concentration was determined from its $A_{260}$ using a Thermo Scientific NanoDrop 2000c and the concentration was adjusted to 50 ng/µL prior to storage at -20°C.

### 2.3.4 Oligonucleotide Design and Preparation

Mutagenic oligonucleotide primers were individually designed SDM as described within the guidelines of the QuikChange XL Site-Directed Mutagenesis Kit Instruction Manual (Agilent Technologies n.d.).

Oligonucleotide primers were synthesized by Invitrogen™ and were diluted to 50 ng/µL with the addition of nuclease-free $H_2O$, prior to storage at -20°C. The primer sequences used for SDM are listed in the Appendix.

### 2.3.5 Site-Directed Mutagenesis

All mutagenesis reactions were completed using the protocol supplied with the Quikchange® II XL Site-Directed Mutagenesis Kit by Agilent Technologies. In brief, single amino acid substitutions were introduced into the parental double-stranded DNA vector using *PfuUltra* HF DNA polymerase for strand synthesis, this was followed by *Dpn* I endonuclease restriction digestion (target sequence: 5'-Gm$^6$ATC-3') of the parental template. The undigested vector containing the desired mutation was then transformed into the provided XL10-Gold ultracompetent cell strain for nick repair.

All kit components, plasmids and oligonucleotide primers were thawed on ice. PCR reactions were prepared as recommended by Agilent Technologies and placed in Fisherbrand 0.5 mL thin-walled PCR tubes. Amplification cycling parameters (Table 2.4) were followed as suggested (Agilent Technologies n.d.) using a Techne TC-512 PCR

thermal cycler. *Dpn* I digestion immediately followed strand synthesis, with incubation of reactions at 37°C in a water bath for 1 hour, prior to transformation into XL10-Gold ultracompetent cells.

Table 2.4: Cycling parameters for the Quikchange® II XL method

| Segment | Cycles | Temperature (°C) | Time |
|---|---|---|---|
| 1 | 1 | 95 | 1 minute |
| 2 | 18 | 95 | 50 seconds |
| | | 60 | 50 seconds |
| | | 68 | 1 minute/kb of plasmid length |
| 3 | 1 minute | 68 | 7 minutes |

The pWhitescript mutagenesis control and the pUC18 transformation control were used at all times. Mutagenic efficiency was determined from the number of colonies displaying a blue phenotype indicative of acquisition of β-galactosidase production.

### 2.3.6 Plasmid DNA sequencing

Sequencing of all recombinant plasmid DNA constructs was completed using Applied Biosystems BigDye Terminator chemistry on an ABI PRISM 3730xl DNA Analyzer, as part of the QuickLane Express Sequencing service provided by Beckman Coulter Genomics. DNA sequence analysis was visualized and performed using FinchTV sequencing chromatogram trace viewer V1.4, provided by Geospiza.

### 2.3.7 IPTG-induced heterologous protein expression test in *E. coli* expression strains

*E. coli* strains BL21 (DE3) and BL21-CodonPlus (DE3)-RIPL were used for protein expression. Seven colonies were selected from the transformation plate and added to 30 mL Sterilin tubes containing 10 mL of liquid LB media supplemented with appropriate antibiotic prior to incubation with shaking (250 rpm) at 37°C for 12-16 hours. From these starter cultures (10 µL) were used to inoculate (1:1000) 30 mL Sterilin tubes

containing 10 mL of liquid LB media supplemented with appropriate antibiotic. These cultures were then grown for approximately 3 hours with shaking (250 rpm) at 37°C until an $OD_{600}$ of 0.4-0.6 was reached, at which point 1 mL from each was taken as a pre-induction sample and the $OD_{600}$ recorded. These samples were then centrifuged in separate 1.5 mL eppendorf tubes for 5 minutes at 3000 rpm and the pellets retained and stored at -20°C for subsequent sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) analysis. Remaining samples were induced with a final concentration of 0.4 mM IPTG and incubated for an additional 3 hours at 37°C with shaking (250 rpm) prior to recording the $OD_{600}$ for comparison. 1 mL was taken as a post-induction samples and centrifuged at 3000 rpm for 5 minutes. Pellets were retained and stored at -20°C.

Thawed pellets were resuspended in 20 µL of SDS-PAGE loading buffer and heated at 95°C for 10 minutes. Protein expression levels were determined by SDS-PAGE analysis (Section 2.3.19) of pre- and post-induced samples, based on comparable cell density.

Densitometric analysis of the SDS-PAGE gels was used to identify high heterologous protein expressing colonies suitable for creating glycerol stocks. 1 mL of the relevant starter culture was added to a 1.5 mL Eppendorf tube and centrifuged at 3000 rpm for 5 minutes to harvest cells. The pellet was then resuspended in 1 mL of 30% (v/v) glycerol prior to storage at -80°C.

### 2.3.8 IPTG-induced protein expression in *E. coli* expression strains using nutrient-rich media

10 mL of liquid LB media supplemented with appropriate antibiotic was added to a 30 mL Sterilin tube. This was inoculated from stored *E. coli* protein expression glycerol stocks prior to overnight incubation with shaking (250 rpm) at 37°C.

0.5 mL of the overnight starter culture was added to 500 mL of LB antibiotic supplemented liquid media. Cells were grown at 37°C with shaking at 250 rpm until an $OD_{600}$ of 0.6 was reached. Protein expression was induced with 0.4 mM IPTG and cultures were incubated as described in the protein-specific method. 1 mL was taken from pre- and post-induction samples and protein expression was analysed as previously described (Section 2.3.7).

Cells were harvested in 1 L polypropylene bottles by centrifugation at 7000 rpm for 20 minutes at 4°C in a Beckman Coulter Avanti J-20XP centrifuge equipped with a

JLA 8.1000 rotor. Pellets were collected and routinely stored at -20°C for up to 12 months.

### 2.3.9 IPTG-induced heterologous protein expression in *E.coli* expression strains using isotopically enriched minimal media

M9 salts minimal media was made (components per 975 mL: 6 g $Na_2HPO_4$, 3 g $KH_2PO_4$, 0.5 g NaCl) and autoclaved at 120°C for 20 minutes at 15 psi. *E. coli* trace elements were made (components per L: 5 g $Na_2EDTA$ dissolved in 800 mL $ddH_2O$, with addition of 5 M HCl adjusting to pH 7 after adding each of the following: 0.5 g $FeCl_3$, 0.05 g $ZnCl_2$, 0.01 g $CuCl_2.2H_2O$, 0.01 g $CoCl_2.6H_2O$, 0.01 g $H_3BO_3$), filter-sterilised (0.22 µm) and stored at room temperature.

Heat-sensitive components were filter-sterilised (0.22 µm) and added to the autoclaved M9 salts minimal media (5 mL 20% (w/v) $^{15}(NH_4)_2 SO_4$, 10 mL 20% (w/v) D-Glucose, 1 mL 1 $MgSO_4$, 1 mL 0.1M $CaCl_2$, 1 mL 1M thiamine, 10 mL *E. coli* trace elements and appropriate antibiotics). Then 10 mL of the isotopically enriched complete minimal media was decanted into a 30 mL Sterilin tube and inoculated from stored *E. coli* protein expression glycerol stocks. This starter culture was grown overnight with shaking (250 rpm) at 37°C.

The 10 mL starter culture was reintroduced to the remaining minimal media and incubated at 37°C with shaking at 250 rpm until an $OD_{600}$ of 0.6 was reached. Protein expression was induced with 0.4 mM IPTG and cultures were incubated for a further 4 hours at 37°C with shaking at 250 rpm. 1 mL was taken from pre- and post-induction samples and protein expression was analysed as previously described (Section 2.3.7).

Cells were harvested in 1 L polypropylene bottles by centrifugation at 7000 rpm for 20 minutes at 4°C in a Beckman Coulter Avanti$^{TM}$ J-20XP centrifuge equipped with a JLA 8.1000 rotor. Pellets were collected and routinely stored at -20°C for up to 12 months.

### 2.3.10 Physical disruption of *E. coli* cells by sonication

Cell pellets, previously obtained from 1 L of culture, were thawed on ice, and resuspended in 40-50 mL of appropriate buffer with the addition of one cOmplete, EDTA-free Protease Inhibitor Cocktail tablet (Roche). Bacterial cells were disrupted by

sonication using a Sonics Vibra-cell (pulse on: 3.3 s; pulse off: 9.9 s; total pulse on time: 10 minutes; probe temperature: 20°C; amplitude: 30%). The cell suspension was kept on ice at all times.

### 2.3.11 Isolation of soluble proteins

Post-sonicated cell suspension containing expressed soluble protein, was centrifuged at 18000 rpm for 30 minutes at 4 °C in a Beckman Coulter Avanti J-20XP centrifuge (JA-25.50 rotor) to remove cellular debris. Supernatant was retained for further purification by chromatography.

### 2.3.12 Purification and solubilisation of *E. coli* inclusion bodies

Inclusion bodies were purified and solubilised as described in the protein-specific methods. The presence or absence of the target protein was monitored at all stages by SDS-PAGE analysis (Section 2.3.19).

### 2.3.13 Refolding of heterologous proteins by serial dialysis

Concentrations of partially purified protein-containing solutions were assessed at $OD_{280}$ using a Thermo Scientific NanoDrop 2000c. Samples were diluted to protein concentrations of 0.2 mg/mL with appropriate buffers (see protein-specific methods). Using Spectrum Standard Regenerated Cellulose Spectra/Por dialysis membrane MWCO 3.5 kDa (Spectrum Laboratories, Inc.), the protein solution was subjected to repeated overnight dialysis against a series of buffers as described in the protein-specific methods.

### 2.3.14 Purification of heterologous proteins by Immobilised Metal Affinity Chromatography

Samples were filtered (0.22 μm) and applied at 1 mL/min to a pre-equilibrated 5 mL HisTrap FF column prepacked with Ni Sepharose 6 Fast Flow resin using the ÄKTAprime system. Samples exposed to potential protease degradation were purified using the ÄKTApurifier system at 4°C. Unbound proteins were collected in 4 mL fractions. Bound proteins were eluted with a 0-100% gradient of appropriate buffer supplemented with 500 mM imidazole and collected in 1 mL fractions for further SDS-

PAGE analysis (Section 2.3.19).

### 2.3.15 Purification of heterologous proteins by Cation Exchange Chromatography

Refolded protein solutions were filtered (0.22 μm) and thoroughly degassed prior to application at 10 mL/min to a 30 mL SP-Sepharose FF XK26 column using the ÄKTAprime system (GE Healthcare Life Sciences) that had been pre-equilibrated with an appropriate low salt buffer (see protein-specific methods). Bound protein of interest was eluted with a 0-100% gradient of appropriate high salt buffer and 1 mL fractions were collected for SDS-PAGE analysis (Section 2.3.19).

### 2.3.16 N-terminal $His_6$-tag removal

Samples were filtered (0.22 μm) and applied at 1 mL/min to a HisTrap FF column prepacked with Ni Sepharose 6 Fast Flow resin and pre-equilibrated with an appropriate buffer using the ÄKTAprime system. Unbound proteins were washed from the column with 5 column volumes prior to column detachment. Thrombin (25 U) was diluted in 5 mL appropriate buffer and manually injected onto the column. The column was sealed and incubated overnight in a 37 $^{o}$C oven. A 5.4 mL p-amidobenzamidine column was attached for thrombin removal and cleaved proteins were eluted with 5 column volumes of appropriate buffer at 2.5 mL/min and collected in 2 mL fractions. Bound proteins were then eluted with a 0-100% gradient of an appropriate buffer supplemented with 500 mM imidazole and collected in 2 mL fractions for further SDS-PAGE analysis (Section 2.3.19).

### 2.3.17 Purification of heterologous proteins by Anion Exchange Chromatography

Protein solutions were filtered (0.22 μm) and thoroughly degassed prior to application at 5 mL/min to a 30 mL Q-Sepharose FF XK26 column using the ÄKTAprime system that had been pre-equilibrated with an appropriate low salt buffer (see protein-specific methods). Bound protein of interest was eluted with a 0-100% gradient of appropriate high salt buffer and 1 mL fractions were collected for SDS-PAGE analysis (Section 2.3.19).

### 2.3.18 Preparation of Tris-Tricine SDS-PAGE gels and reagents

Using Novex 1.0 mm plastic gel cassettes (Life Technologies), 16.5% Tris-Tricine SDS-PAGE gels were cast. A 30 mL volume, suitable for 4 cassettes was made accordingly - 16.5 mL Protogel (30% w/v) 37.5:1 Acrylamide:Bisacrylamide stabilised solution, 10 mL 3X Tris-Tricine gel buffer, 3.1 mL glycerol, 0.4 mL distilled $H_2O$, 100 µL 10% (w/v) ammonium persulphate (APS), 40 µL N, N, N′, N′-Tetramethylethylenediamine (TEMED). For 12% Tris-Tricine SDS-PAGE gels the volumes of Protogel (37.5:1) solution and $H_2O$ was substituted with 12 mL and 4.9 mL, respectively. The separating gel was overlayed with distilled $H_2O$, and allowed to polymerise for 1 hour before removal of the $H_2O$ and addition of a 6% stacking gel (volume suitable for 4 cassettes – 1.5 mL Protogel 37.5:1, 2.5 mL 3X Tris-Tricine gel buffer, 6 mL distilled $H_2O$, 50 µL 10% (w/v) APS, 20 µL TEMED) and appropriate sized comb. Cast SDS-PAGE gels were inverted in distilled $H_2O$ and stored at 4°C.

3X Tris-Tricine gel buffer (3 M Tris-HCl pH 8.45, 0.3% (w/v) SDS) was stored at room temperature along with Anode buffer (100 mM Tris-HCl pH 8.9) and Cathode buffer (100 mM Tris-HCl pH 8.25, 100 mM Tricine, 0.1% (w/v) SDS).

A Coomassie Blue stain solution (50% Methanol, 10% Acetic acid, 0.25% (w/v) Coomassie Brilliant Blue R-250 dye) was used for all SDS-PAGE protein staining.

### 2.3.19 Analysis of protein samples by Tris-Tricine SDS-PAGE

Appropriate sample volumes were added to 2X SDS-PAGE reducing loading buffer (126 mM Tris-HCl pH 6.8, 20% (w/v) glycerol, 4% (w/v) SDS, 0.005% (w/v) bromophenol blue, 5% (v/v) β-mercaptoethanol) and heated at 95°C for 10 minutes. Samples were loaded onto 16.5% Tris Tricine SDS-PAGE gels (Section 2.3.18) alongside 5 µL of SeeBlue Plus2, a pre-stained molecular weight standard, and were run for 20 minute at 80 V followed by 45-55 minutes at 180 V unless otherwise stated.

### 2.3.20 Screening proMMP-1* activation by limited proteolysis

#### 2.3.20.1 Trypsin

Trypsin (Sigma Aldrich) was freshly dissolved in TNC pH 7.4 (50 mM Tris-HCl pH 7.4, 150 mM NaCl, 5 mM $CaCl_2$ and 0.02% (w/v) $NaN_3$) and serially diluted prior to incubation with 10 µM of proMMP-1*. Samples were then incubated at 37 °C for 5

minutes. Reactions were stopped with 1 mM phenylmethylsulfonyl fluoride (PMSF) and heated at 95 °C for 5 minutes prior to analysis by SDS-PAGE ( Section 2.3.19).

#### 2.3.20.2 CAT-3

An equimolar concentration of recombinantly produced CAT-3 was added to 10 µM proMMP-1* in TNC pH 7.4. Samples were then incubated at 37 °C for 0-24 hours. Reactions were stopped with 2X SDS-PAGE reducing loading buffer supplemented with 20 mM EDTA and heated at 95 °C for 5 minutes prior to analysis by SDS-PAGE ( Section 2.3.19).

### 2.3.21 APMA-induced activation of proMMPs

Purified proteins were incubated with 0.5 mM APMA for an appropriate time (see protein-specific methods) in a 37°C waterbath. APMA was removed using a 35 mL HiPrep desalting column prepacked with Sephadex G-25 Fine resin pre-equilibrated with an appropriate buffer using the ÄKTAprime system. Samples were manually injected directly onto the column using a glass luer-lock syringe (Fisher Scientific) and eluted using a flow rate of 3 mL/min. APMA was then disposed of according to local regulations.

### 2.3.22 Purification of heterologous proteins by Gel Filtration Chromatography

Protein-containing samples were concentrated, using centrifugal filter units with an appropriate MWCO (Millipore), to a volume of ~ 0.5-2 mL using a Beckman Coulter Allegra 25R refrigerated centrifuge (rotor AT-14-10) pre-chilled to 4 °C. Samples were resuspended after each spin (3000 rpm for 1-10 minutes) and applied at 1 mL/min to a HiLoad Superdex 75 26/60 prep grade column pre-equilibrated with an appropriate buffer (see protein-specific methods). 5 mL fractions were collected and analysed by SDS-PAGE (Section 2.3.19).

### 2.3.23 Storage of heterologous proteins

Both zymogen and mature proteins were routinely stored at -80°C diluted 50:50 in 2x Storage Buffer (50 mM Tris-HCl, pH 7.5, 450 mM NaCl, 40 µM $ZnCl_2$, 1% (w/v Brij 35, 60% glycerol) for up to 6 months.

### 2.3.24 Continuous fluorogenic assay of CAT-3

Lyophilised quenched fluorogenic MMP-3 substrate (Mca-Arg-Pro-Lys-Pro-Val-Glu-Nva-Trp-Arg-Lys-(Dnp)-$NH_2$ (Nagase, Fields et al. 1994) was purchased from Bachem and prepared as 100 µM stock solutions in DMSO. The substrate is labelled with a fluorescent donor (7-methoxycoumarin-4-yl)acetyl (Mca) and quenching moiety 2,4-dinitrophenyl (Dnp), which flank the MMP-3 susceptible bond (Glu-Nva). Cleavage of this bond results in separation of the fluorophore and quencher concomitant with increased fluorescence emission. Fluorescent assays were performed at $\lambda_{ex}$ 325 nm and $\lambda_{em}$ 393 nm using a Perkin Elmer LS 50B fluorescence spectrometer. Typically, hydrolysis assays were conducted at ambient temperature by incubating 10 nM of CAT-3 with 100 nM of substrate in TNC pH 7.4 to assess CAT-3 catalytic activity.

### 2.3.25 Continuous spectrophotometric assay of MMP-1 and CAT-1

Lyophilised chromogenic MMP substrate (Ac-Pro-Leu-Gly-[2-mercapto-4-methyl-pentanoyl]-Leu-Gly-$OC_2H_5$) was purchased, and prepared as 5 mM stock solutions in DMSO and stored at -20 °C in 1.5 mL black polypropylene LiteSafe microcentrifuge tubes (Sigma Aldrich). Hydrolysis of the thioester substrate bond by MMP-1 produces a free sulfhydryl group, which reacts with 5,5-dithio-bis(2-nitrobenzoic acid) (DTNB; Ellmans' reagent) to form 5-mercapto-2-nitrobenzoic acid (TNB). Activity is detected indirectly by measuring TNB formation, which absorbs at 412 nm. Assays were conducted at 37 °C by incubating 0-10 nM of enzyme with 25 µM of substrate in assay buffer (50 mM HEPES pH 7, 10 mM $CaCl_2$, 0.05% Brij-35 and 1 mM DTNB) and monitored continuously at $\lambda_{412}$ for 10 minutes using a Thermo Scientific NanoDrop 2000c. The absorbance data, taken for 3 minutes after the initial 30 seconds, were fitted by linear regression using Grafit Version 5 (Erithacus Software Ltd).

### 2.3.26 Type I collagen preparation

Lyophilised bovine type I collagen (Sigma Aldrich) was reconstituted with 0.5 M acetic acid to yield a 1 mg/mL solution, which was then solubilised overnight with gentle stirring at 4 $^{o}$C. Using 10 kDa MWCO Spectrum Standard Regenerated Cellulose Spectra/Por dialysis membrane (Spectrum Laboratories, Inc.), the collagen solution was subjected to overnight dialysis at 4 $^{o}$C against 1000 volumes of TNC pH 7.4.

### 2.3.27 Collagenase cleavage of Type I collagen

Previously prepared collagen solutions were diluted to 0.1 mg/mL with TNC pH 7.4 and incubated with a 30 fold molar excess of MMP-1 generated by APMA initiated autolysis (Section 2.3.21). Reactions were then incubated at 37 $^{o}$C for 20 hours prior to the addition of SDS-PAGE loading buffer supplemented with 20 mM EDTA to stop further MMP-1 cleavage activity. Products were analysed by SDS-PAGE (Section 2.3.19) and examined for production of ¾ and ¼ collagen fragments.

### 2.3.28 Nuclear Magnetic Resonance (NMR) sample preparation

Purified isotopically $^{15}$N-labelled protein was dialysed exhaustively against a suitable buffer and concentrated to 0.2-0.5 mM in a final volume of 300 µL using a centrifugal filter unit with appropriate MWCO (Millipore) as detailed (Section 2.3.22). HPX-1 protein samples were dialysed against 20 mM NaOAc pH 4.8, 5 mM $CaCl_2$, 50 µM $ZnCl_2$ and 0.02% (w/v) $NaN_3$ and proCAT-1* protein samples were dialysed against 10 mM Tris-HCl pH 7.2, 5 mM $CaCl_2$, 150 mM NaCl and 0.02% (w/v) $NaN_3$. 10% $D_2O$ was added to all samples in preparation for deuterium signal locking of the NMR spectrometer and transferred to $D_2O$-matched 5 mm symmetrical NMR Shigemi microtubes (Sigma Aldrich). Samples were stored at 4 $^{o}$C prior to NMR data collection.

### 2.3.29 NMR spectral acquisition, processing and analysis

Experiments were routinely performed on an Agilent Inova 600 MHz spectrophotometer, operating at 14.1 Tesla and equipped with a triple resonance cryoprobe, at the University of Portsmouth. 1D $^{1}$H NMR and 2D $^{1}$H-$^{15}$N heteronuclear single quantum coherence (HSQC) spectra of the target protein was acquired as 4096 and 256 complex points using spectral widths of 9900.99 and 2431.06 Hz for $F_1$ and $F_2$ dimensions, respectively. This gave a resolution of 19 Hz in the

nitrogen direction and 4.83 Hz in the hydrogen direction. The spectral data was acquired and processed by Dr Chris Read, University of Portsmouth, using NMRPipe (Delaglio, Grzesiek et al. 1995) and visualised using CcpNmr Analysis v2.2.2 (Vranken, Boucher et al. 2005). NMR spectral quality was examined and assessed primarily on spectral dispersion, line widths and number of resolved peaks.

Construct-specific methods are described in the subsequent tables (Tables 2.5-2.8) and contain details of both the purification pathway required and conditions necessary for optimal production of the recombinant protein of interest.

Table 2.5: Protein-specific protocol for CAT-3

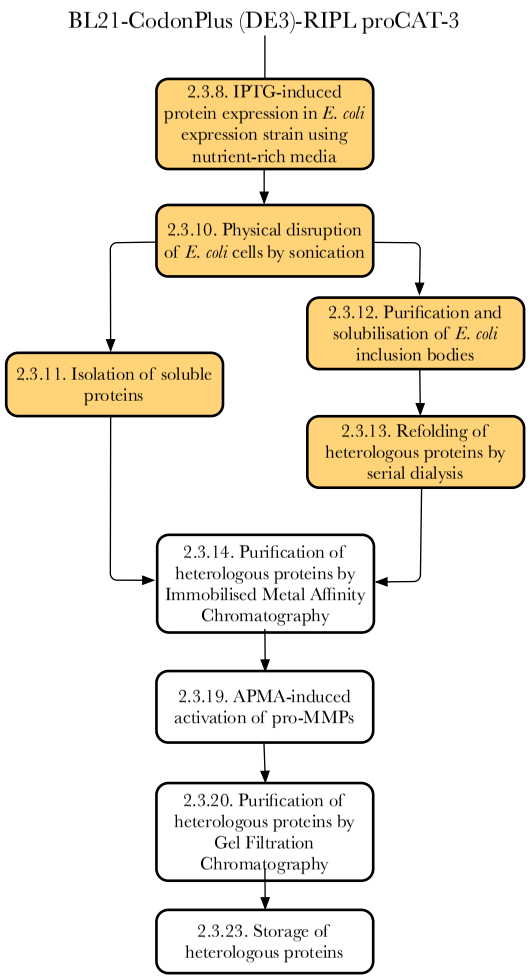
| Purification Method | Method Section | Method Details |
|---|---|---|
|  | 2.3.8: IPTG-induced protein expression in *E. coli* expression strains using nutrient-rich media | Liquid LB media was supplemented with 50 µg/mL kanamycin and 34 µg/mL chloramphenicol. Cells were grown at 37°C with shaking at 250 rpm until an $OD_{600}$ of 0.6 was reached. Protein expression was induced with 0.4 mM IPTG and cultures were incubated overnight at 20°C with shaking at 250 rpm. |
| | 2.3.10: Physical disruption of *E. coli* by sonication | 1 L cell pellets were thawed on ice and resuspended in 40 mL of Day 2 RF prCAT-3 buffer suitable for soluble proteins (Day 2 RF - see below), prior to sonication. |
| | 2.3.11: Isolation of soluble proteins | Post-sonicated cell suspension containing expressed soluble protein, was centrifuged at 18000 rpm for 30 minutes at 4 ˚C. |
| | 2.3.12: Purification and solubilisation of *E. coli* inclusion bodies | Pellet was resuspended in 10 mL of 8 M urea and resonicated as described previously (Section 2.3.10). Solubilised inclusion bodies were centrifuged at 18000 rpm for 20 minutes at 4 ˚C and protein-containing supernatant was retained. |
| | 2.3.13: Refolding of heterologous proteins by serial dialysis | Insoluble protein supernatant was diluted to 0.2 mg/mL using 8 M urea and subjected to serial overnight dialysis:<br>**Day 1**: 10 volumes of 50 mM Tris-HCl pH 8.2,10 mM $CaCl_2$, 0.1 mM $ZnCl_2$, 5 mM DTT, 0.1% (w/v) Brij 35, 450 mM sucrose and 0.02% (w/v) $NaN_3$.<br>**Day 2**: 50 mM Tris-HCl pH 8.2, 10 mM $CaCl_2$, 0.2 mM $ZnCl_2$, 150 mM NaCl, 0.1% (w/v) Brij 35, 450 mM sucrose and 0.02% (w/v) $NaN_3$. |

Table 2.5 (cont.): Protein-specific protocol for CAT-3

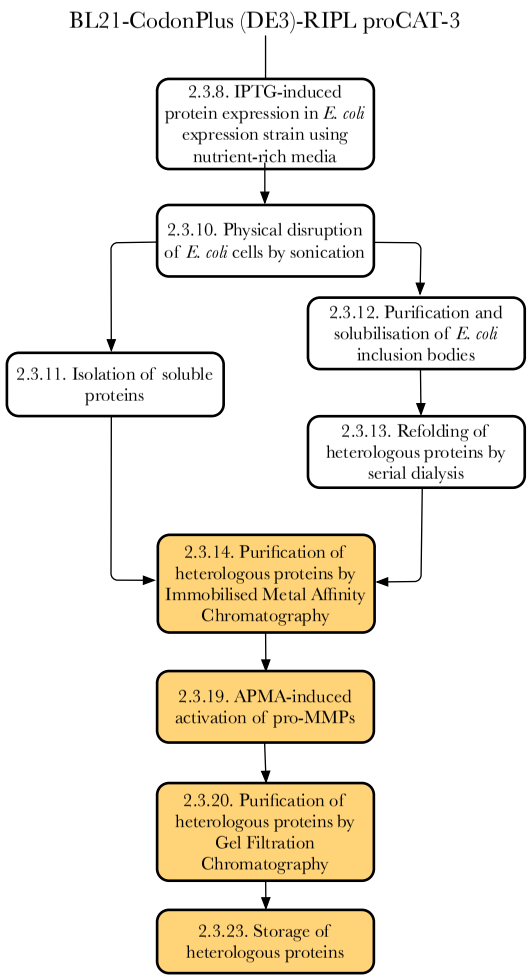
| Purification Method | Method Section | Method Details |
|---|---|---|
|  | 2.3.14: Purification of heterologous proteins by Immobilised Metal Affinity Chromatography | Samples were applied to a HisTrap FF column pre-equilibrated with Day 2 RF buffer. Bound protein was eluted with 0-100% gradient of Day 2 RF buffer B (supplemented with 500 mM Imidazole). |
| | 2.3.21: APMA-induced activation of pro-MMPs | Purified proteins were dialysed against 1000 volumes of TNC pH 7.4. Samples were then incubated with 0.5 mM APMA for 6 hours in a 37°C waterbath, prior to APMA removal. |
| | 2.3.22: Purification of heterologous proteins by Gel Filtration Chromatography | Concentrated protein samples were applied to a HiLoad Superdex 75 26/60 prep grade column pre-equilibrated with TNC pH 7 .4. |

Table 2.6: Protein-specific protocol for HPX-1

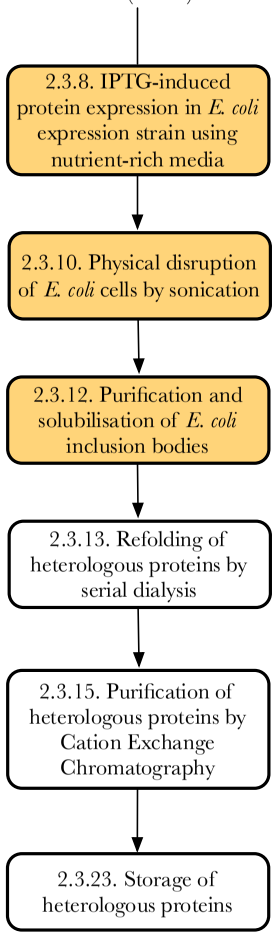
| Purification Method | Method Section | Method Details |
|---|---|---|
| BL21-CodonPlus (DE3)-RIPL HPX-1<br>  | 2.3.8: IPTG-induced protein expression in *E. coli* expression strains using nutrient-rich media | Liquid LB media was supplemented with 100 µg/mL ampicillin and 34 µg/mL chloramphenicol. Cells were grown at 37°C with shaking at 250 rpm until an $OD_{600}$ of 0.6 was reached. Protein expression was induced with 0.4 mM IPTG and cultures were incubated for 4 hours at 37°C with shaking at 250 rpm. |
| | 2.3.10: Physical disruption of *E. coli* by sonication | 1 L cell pellets were thawed on ice and resuspended in 50 mL of Sonication/Lysis Buffer (50 mM Tris-HCl, pH 8.0, 1 M NaCl and 1 mM EDTA), prior to sonication. |
| | 2.3.12: Purification and solubilisation of *E. coli* inclusion bodies | 1% (w/v) deoxycholic acid was added to crude lysate and stirred for 1 hour at room temperature. 200 µL of DNase I (28.4 U/mL) was added with additional stirring for 1 hour at room temperature. Cellular debris was removed by centrifugation at 14000 rpm for 20 minutes at 4 ˚C. The resulting pellet was resuspended in 20 mL of Sonication/Lysis Buffer and stirred for 30 minutes at ambient temperature. Inclusion bodies were isolated by centrifugation as described above and the pellet was resuspended in 20 mL of Inclusion Body Pellet Wash Buffer (50 mM Tris pH 8.5, 0.1 M NaCl, 1 mM EDTA and 0.5% (w/v) Triton X100). A further centrifugation and 30 minute wash step with Lysis Buffer was conducted. The final pellet was resuspended in 20 mL of Solubilisation Buffer (8 M urea, 0.25 M GdmCl, 50 µM $ZnCl_2$, 20 mM Tris pH 8.6) with the addition of 20 mM fresh DTT and stirred overnight at 4°C. The solubilised inclusion body cell suspension was centrifuged at 14000 rpm for 20 minutes at 4 ˚C and the supernatant was retained. |

Table 2.6 (cont.): Protein-specific protocol for HPX-1

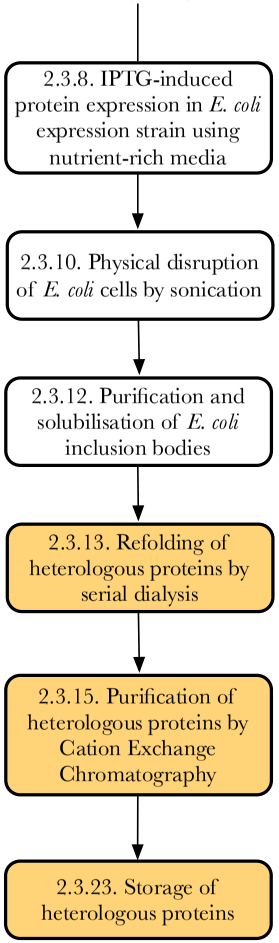
| Purification Method | Method Section | Method Details |
|---|---|---|
| BL21-CodonPlus (DE3)-RIPL HPX-1<br>  <br> | 2.3.13: Refolding of heterologous proteins by serial dialysis | Insoluble protein supernatant was diluted to 0.2 mg/mL using 50 mM Tris-HCl pH 8.0, 6 M urea, 10 mM $CaCl_2$, 0.1 mM $ZnCl_2$, 0.02% (w/v) $NaN_3$, 20 mM cystamine dihydrochloride, 20 mM DTT and subjected to serial overnight dialysis:<br>**Day 1**: 10 volumes of 50 mM Tris-HCl pH 8.0, 4 M urea, 10 mM $CaCl_2$, 0.1 mM $ZnCl_2$, 0.02% (w/v) $NaN_3$, 5 mM β-ME and 1 mM 2-hydroxyethyl-disulfide.<br>**Day 2**: 10 volumes of 50 mM Tris-HCl pH 7.2, 2 M urea, 10 mM $CaCl_2$, 0.1 mM $ZnCl_2$, 0.15 M NaCl and 0.02% (w/v) $NaN_3$.<br>**Day 3**: 10 volumes of 50 mM Tris-HCl, pH 7.2, 10 mM $CaCl_2$, 0.1 mM $ZnCl_2$, 0.15 M NaCl, 0.02% (w/v) $NaN_3$. |
| | 2.3.15: Purification of heterologous proteins by Cation Exchange Chromatography | Refolded samples were subsequently dialysed against 10 volumes of CEX A Buffer: 20 mM NaOAc pH 4.8, 5 mM $CaCl_2$, 50 µM $ZnCl_2$, 0.02% (w/v) $NaN_3$.<br>Refolded protein solutions were applied to a 30 mL SP-Sepharose FF XK26 column that had been pre-equilibrated with CEX Buffer A. Bound protein was eluted with 0-100% gradient of CEX Buffer B (20 mM NaOAc pH 4.8, 5mM $CaCl_2$, 50 µM $ZnCl_2$, 1 M NaCl and 0.02% (w/v) $NaN_3$. |

Table 2.7: Protein-specific protocol for proMMP-1

| Purification Method | Method Section | Method Details |
|---|---|---|
| BL21-CodonPlus (DE3)-RIPL proMMP-1 | 2.3.8: IPTG-induced protein expression in *E. coli* expression strains using nutrient-rich media | Liquid LB media was supplemented with 100 µg/mL ampicillin and 34 µg/mL chloramphenicol. Cells were grown at 37°C with shaking at 250 rpm until an $OD_{600}$ of 0.6 was reached. Protein expression was induced with 0.4 mM IPTG and cultures were incubated overnight at 20°C with shaking at 250 rpm. |
| | 2.3.10: Physical disruption of *E. coli* by sonication | 1 L cell pellets were thawed on ice and resuspended in 50 mL of Sonication/Lysis Buffer (50 mM Tris-HCl, pH 8 0, 1 M NaCl and 1 mM EDTA), prior to sonication. |
| | 2.3.12: Purification and solubilisation of *E. coli* inclusion bodies | 1% (w/v) deoxycholic acid was added to crude lysate and stirred for 1 hour at room temperature. 200 µL of DNase I (28.4 U/mL) was added with additional stirring for 1 hour at room temperature. Cellular debris was removed by centrifugation at 14000 rpm for 20 minutes at 4 ˚C. The resulting pellet was resuspended in 20 mL of Sonication/Lysis Buffer and stirred for 30 minutes at ambient temperature. Inclusion bodies were isolated by centrifugation as described above and the pellet was resuspended in 20 mL of Inclusion Body Pellet Wash Buffer (50 mM Tris pH 8.5, 0.1 M NaCl, 1 mM EDTA and 0.5% (w/v) Triton X100). A further centrifugation and 30 minute wash step with Lysis Buffer was conducted. The final pellet was resuspended in 20 mL of Solubilisation Buffer (8 M urea, 0.25 M GdmCl, 50 µM $ZnCl_2$, 20 mM Tris pH 8.6) with the addition of 20 mM fresh DTT and stirred overnight at 4°C. The solubilised inclusion body cell suspension was centrifuged at 14000 rpm for 20 minutes at 4 ˚C and the supernatant was retained. |

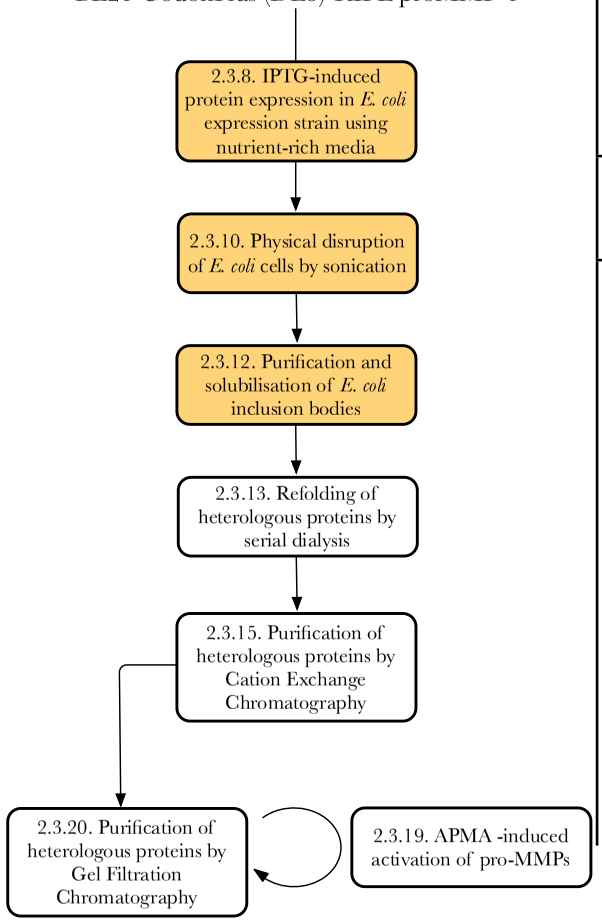


Table 2.7 (cont.): Protein-specific protocol for proMMP-1

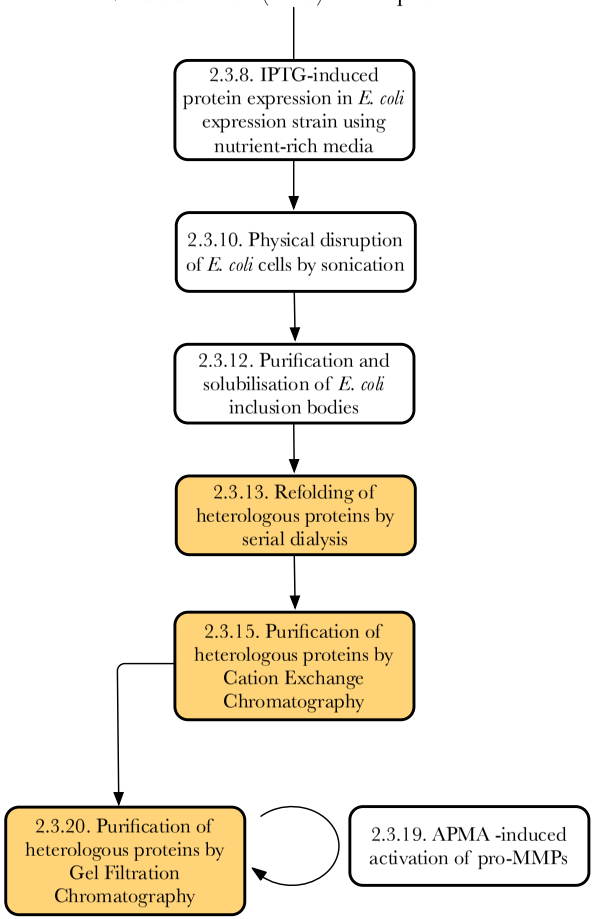
| Purification Method | Method Section | Method Details |
|---|---|---|
| BL21-CodonPlus (DE3)-RIPL proMMP-1<br>  <br> | 2.3.13: Refolding of heterologous proteins by serial dialysis | Insoluble protein supernatant was diluted to 0.2 mg/mL using 50 mM Tris-HCl pH 8.0, 6 M urea, 10 mM $CaCl_2$, 0.1 mM $ZnCl_2$, 0.02% (w/v) $NaN_3$, 20 mM cystamine dihydrochloride, 20 mM DTT and subjected to serial overnight dialysis:<br>**Day 1**: 10 volumes of 50 mM Tris-HCl pH 8.0, 4 M urea, 10 mM $CaCl_2$, 0.1 mM $ZnCl_2$, 0.02% (w/v) $NaN_3$, 5 mM β-ME and 1 mM 2-hydroxyethyl-disulfide.<br>**Day 2**: 10 volumes of 50 mM Tris-HCl pH 7.2, 2 M urea, 10 mM $CaCl_2$, 0.1 mM $ZnCl_2$, 0.15 M NaCl and 0.02% (w/v) $NaN_3$.<br>**Day 3**: 10 volumes of 50 mM Tris-HCl pH 7.2, 10 mM $CaCl_2$, 0.1 mM $ZnCl_2$, 0.15 M NaCl, 0.02% (w/v) $NaN_3$. |
| | 2.3.15: Purification of heterologous proteins by Cation Exchange Chromatography | Refolded samples were subsequently dialysed against 10 volumes of CEX A Buffer: 20 mM MES pH 6.5, 150 mM NaCl, 5 mM $CaCl_2$, 50 µM $ZnCl_2$, 0.02% (w/v) $NaN_3$.<br>Refolded protein solutions were applied to a 30 mL SP-Sepharose FF XK26 column that had been pre-equilibrated with CEX Buffer A. Bound protein was eluted with 100% gradient of CEX Buffer B (20 mM MES pH 6.5, 5mM $CaCl_2$, 50 µM $ZnCl_2$, 1 M NaCl and 0.02% (w/v) $NaN_3$). |
| | 2.3.22: Purification of heterologous proteins by Gel Filtration Chromatography | Concentrated protein samples were applied to a HiLoad Superdex 75 26/60 prep grade column pre-equilibrated with TNC pH 7.4. |

Table 2.7 (cont.): Protein-specific protocol for proMMP-1

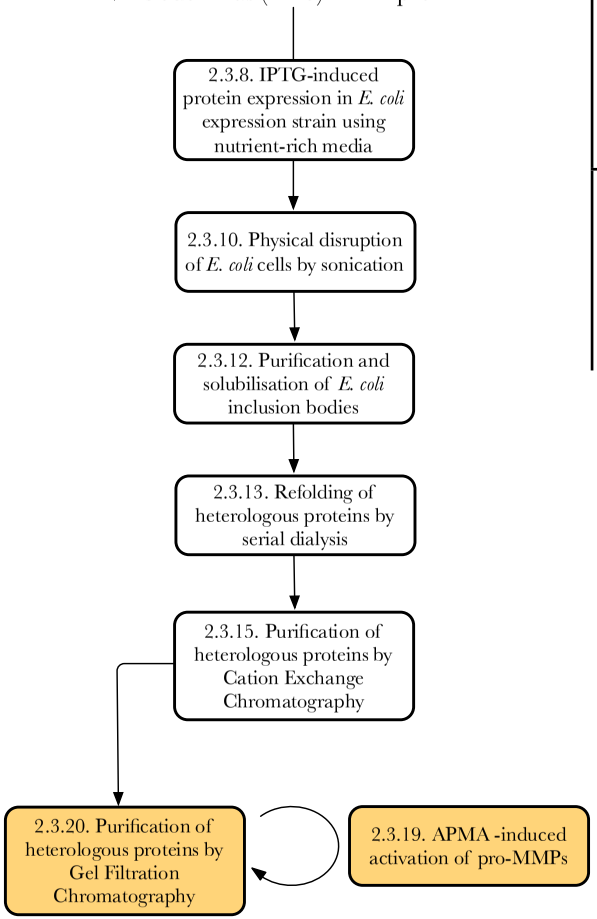
| Purification Method | Method Section | Method Details |
|---|---|---|
| BL21-CodonPlus (DE3)-RIPL proMMP-1<br>  <br> | 2.3.21: APMA-induced activation of pro-MMPs | Purified proteins were dialysed against 1000 volumes of TNC pH 7.4. Samples were then incubated with 0.5 mM APMA for 4 hours in a 37°C waterbath, prior to APMA removal. N.B. For proMMP-1* activation the addition of CAT-3 at equimolar concentrations was necessary. |
| | 2.3.22: Purification of heterologous proteins by Gel Filtration Chromatography | Activated samples were concentrated and reapplied to a HiLoad Superdex 75 26/60 prep grade column pre-equilibrated with TNC pH 7.4. |

Table 2.8: Protein-specific protocol for proCAT-1

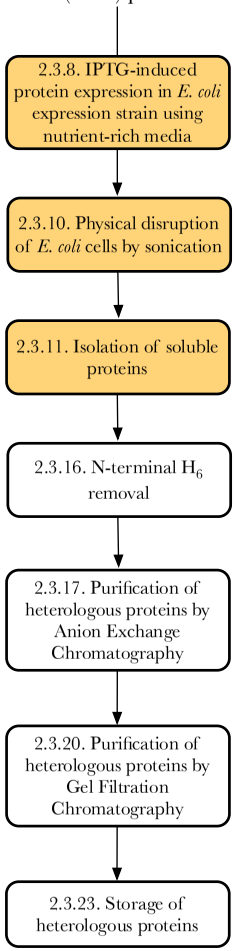
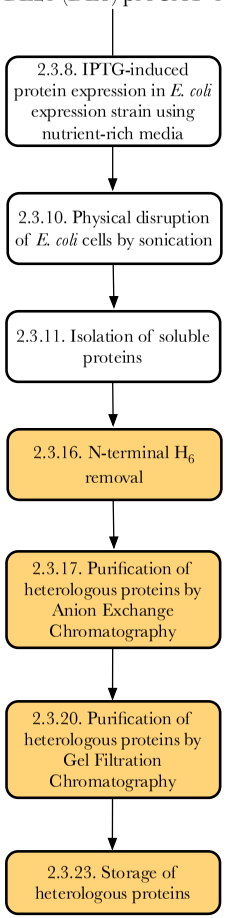
| Purification Method | Method Section | Method Details |
|---|---|---|
|  | 2.3.8: IPTG-induced protein expression in *E. coli* expression strain using nutrient-rich media | Liquid LB media was supplemented with 50 µg/mL kanamycin and 34 µg/mL chloramphenicol. Cells were grown at 37°C with shaking at 250 rpm until an $OD_{600}$ of 0.6 was reached. Protein expression was induced with 0.4 mM IPTG and cultures were incubated overnight at 20°C with shaking at 250 rpm. |
| | 2.310: Physical disruption of *E. coli* cells by sonication | 1 L cell pellets were thawed on ice and resuspended in 40 mL of Day 2 RF buffer suitable for soluble proteins (50 mM Tris-HCl pH 8.2, 10 mM $CaCl_2$, 0.1 mM $ZnCl_2$, 150 mM NaCl, 0.1% (w/v) Brij 35, 450 mM sucrose, 0.02% (w/v) $NaN_3$), prior to sonication. |
| | 2.3.11: Isolation of soluble proteins | Post-sonicated cell suspension containing expressed soluble protein, was centrifuged at 18000 rpm for 30 minutes at 4 ˚C. Supernatant was retained. |

Table 2.8 (cont.): Protein-specific protocol for proCAT-1

| Purification Method | Method Section | Method Details |
|---|---|---|
|  | 2.3.16: N-terminal $H_6$-tag removal. | Samples were filtered (0.22 µm) and applied at 1 mL/min to a HisTrap FF column prepacked with Ni Sepharose 6 Fast Flow resin and pre-equilibrated with Day 2 RF buffer using the ÄKTAprime system. Samples subject to potential protease degradation were purified using the ÄKTApurifier system at 4°C. Unbound proteins were washed from the column with 5 column volumes prior to column detachment. Thrombin (25 U) was diluted in 5 mL of Day 2 RF buffer and manually injected onto the column. The column was sealed and incubated overnight in a 37 °C oven. A 5.4 mL p-amidobenzamidine column was attached for thrombin removal and cleaved proteins were eluted with 5 column volumes of Day 2 RF buffer at 2.5 mL/min and collected in 2 mL fractions. Bound proteins were then eluted with 0-100% gradient of Day 2 RF Buffer B (supplemented with 500 mM imidazole) over 50 mL and collected in 2 mL fractions for further SDS-PAGE analysis. |
| | 2.3.17: Purification of heterologous proteins by Anion Exchange Chromatography | Samples were dialysed overnight against 100 volumes of AEX Buffer A (20 mM Tris-HCl pH 8.6, 5mM $CaCl_2$, 100 µM $ZnCl_2$, 0.02% (w/v) $NaN_3$). Protein solutions were applied to a 30 mL Q-Sepharose XK26 column that had been pre-equilibrated with AEX Buffer A. Bound protein was eluted with 100% gradient of AEX Buffer B (20 mM Tris-HCl pH 8.6, 5mM $CaCl_2$, 100 µM $ZnCl_2$, 1 M NaCl and 0.02% (w/v) $NaN_3$). |
| | 2.3.22: Purification of heterologous proteins by Gel Filtration Chromatography | Samples were concentrated and reapplied to a HiLoad™ Superdex 75 26/60 prep grade column pre-equilibrated with Day 2 RF Buffer A (50 mM Tris-HCl pH 8.2, 10 mM $CaCl_2$, 0.1 mM $ZnCl_2$, 150 mM NaCl, 0.1% (w/v) Brij 35, 450 mM sucrose, 0.02% (w/v) $NaN_3$). |

## 2.4 Results

### 2.4.1 HPX-1 protein

The pET-3a/HPX-1 construct, containing the sequence coding for the human proMMP-1 HPX domain (residues Q253-N450) with an additional methionine residue located at the N-terminus, was used for transformations in both *E. coli* DH5α and BL21-CodonPlus (DE3)-RIPL strains.

Viability of transformed colonies was assessed using small-scale protein expression of randomly selected BL21-CodonPlus (DE3)-RIPL colonies (Section 2.3.7). SDS-PAGE analysis consistently identified significant over-expression of a protein (~ 24 kDa) corresponding to HPX-1 within three hours of treatment with 0.4 mM IPTG (Figure 2.3A). Densitometric analysis of post-induction HPX-1 expression (data not shown) was used to determine which of the transformed strains was most suitable for subsequent large-scale protein production. Appropriate growth conditions had previously been optimised by Dr Laurence Arnold (NIMR, Mill Hill).

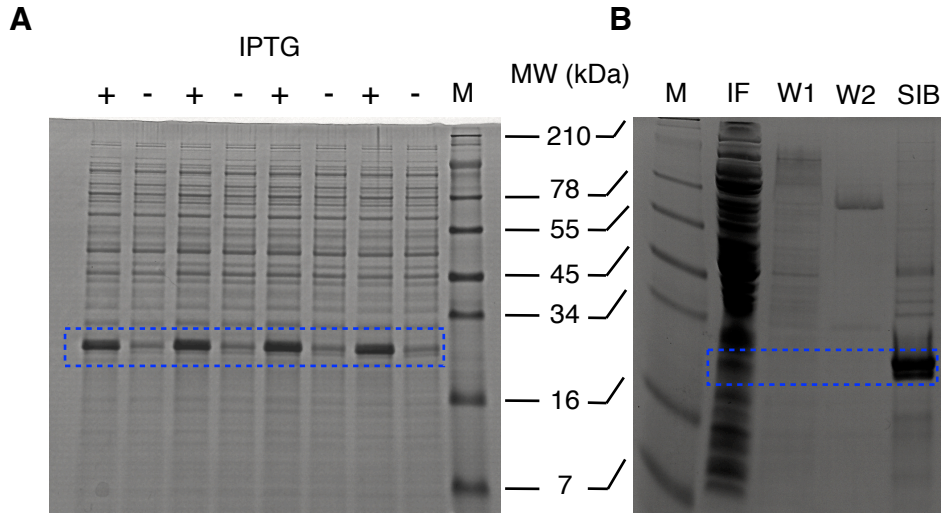



**Figure 2.3: SDS-PAGE analysis of HPX-1 expression and Inclusion Body preparation.**
(**A**) Expression of a protein of ~ 24 kDa (HPX-1) was induced upon treatment with 0.4 mM IPTG and visualised on a 12% (w/v) Tris-Tricine reducing gel, indicated by the *dashed blue box*. (**B**) Insoluble material was harvested by centrifugation (**IF**) and washed (**W1** and **W2**). Pelleted inclusion bodies were then solubilised overnight (**SIB**). All samples were ran alongside SeeBlue Plus2 prestained molecular weight marker (**M**).

As anticipated from a disulphide-containing protein, all HPX-1 protein was found deposited in bacterial inclusion bodies. Centrifugation and washing of this insoluble material produced white crystalline pellets, which after solubilisation (Section 2.3.12) was found to contain almost exclusively (estimated >90% of total protein content)

HPX-1 (Figure 2.3B). An adapted refolding regime (Bertini, Fragai et al. 2009) was then utilised that allowed successful disulphide bond formation between blades 1 and 4. Final purification of the folded HPX-1 domain was carried out using cation exchange chromatography to remove the remaining contaminants (Figure 2.4). This purification strategy yielded ~ 16 mg of >99% homogeneous HPX-1 per 500 mL of LB media (Table 2.9).

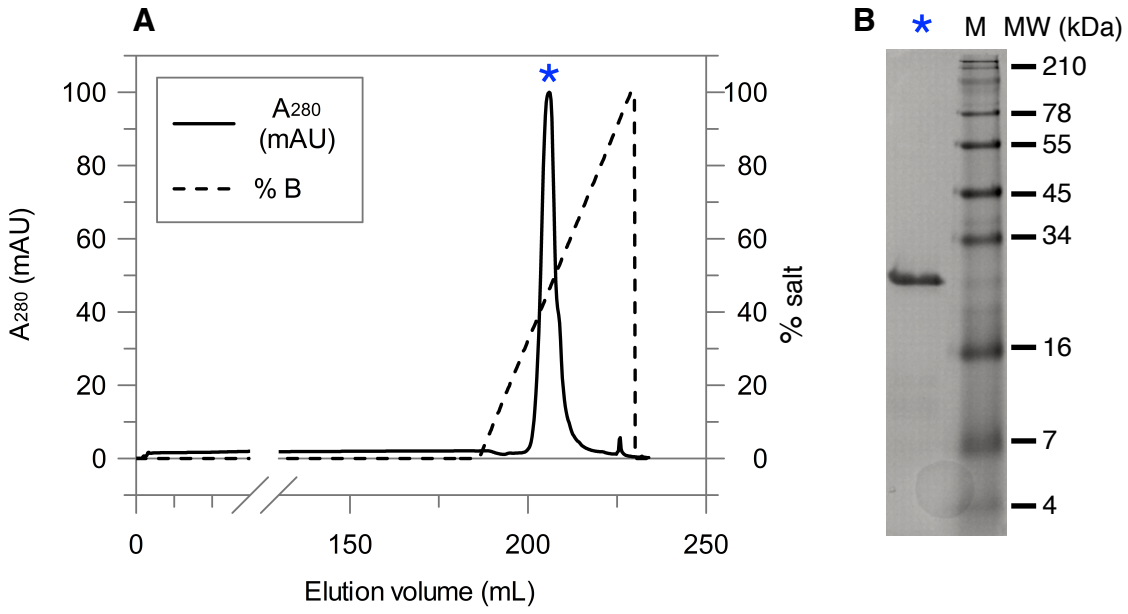



**Figure 2.4: Final purification of HPX-1 using cation exchange chromatography.**
(**A**) Refolded protein solutions were applied to a pre-equilibrated 30 mL SP-Sepharose FF XK26 column at a flow rate of 10 mL/min. Bound protein was eluted with a 0-1 M NaCl gradient. (**B**) Pooled fractions spanning the peak (*) were visualised on a 12% (w/v) Tris-Tricine SDS-PAGE gel. Molecular weight was estimated using SeeBlue® Plus2 prestained molecular weight marker (**M**).

Table 2.9: Purification of MMP proteins[ab]

| Protein | Fraction | Total Volume (mL) | Protein Concentration (mg/mL) | Total Protein (mg) | Recovered Protein (%) |
|---|---|---|---|---|---|
| HPX-1 | Solubilised Inclusion Bodies | 20 | 6.39 | 127.8 | 100 |
| | Post-CEX | 30 | 0.53 | 15.90 | 12.44 |
| proMMP-1* | Solubilised Inclusion Bodies | 20 | 7.87 | 157.4 | 100 |
| | Post-CEX | 30 | 0.44 | 13.2 | 8.38 |
| | Post-GF | 25 | 0.24 | 5.99 | 3.81 |
| proCAT-3 | Solubilised Inclusion Bodies | 20 | 14.02 | 280.40 | 31.53[§] |
| | Soluble protein | 40 | 15.22 | 608.8 | 68.47[§] |
| | Post-IMAC[†] | 17.5 | 1.47 | 25.73 | 4.22 |
| | Post-GF[†] | 3 | 0.67 | 2.02 | 0.33 |
| proCAT-1* | Soluble protein | 40 | 11.3 | 452 | 100 |
| | Post-IMAC | 32 | 1.48 | 47.36 | 10.47 |
| | Post-AEX | 20 | 0.32 | 6.4 | 1.42 |
| | Post-GF | 3.7 | 1.39 | 5.14 | 1.13 |

[a] from 500 mL of culture
[b] protein concentration was determined spectrophotometrically ($A_{280}$) using the molar extinction coefficient values obtained from ProtParam analysis (Gasteiger, Gattiker et al. 2003). Assumes pure protein.
[§] percentage of total expressed protein
† data taken from soluble fraction and is therefore expressed as a percentage of total soluble protein

The conformation of the refolded HPX-1 domain was assessed by acquisition of a 2D $^{1}H$-$^{15}N$ HSQC spectrum (Section 2.3.29). The $^{1}H$-$^{15}N$ HSQC spectrum shown in Figure 2.5 indicates extensive chemical shift dispersion characteristic of a folded protein, and therefore the current refolding strategy is suitable for HPX-1 production.

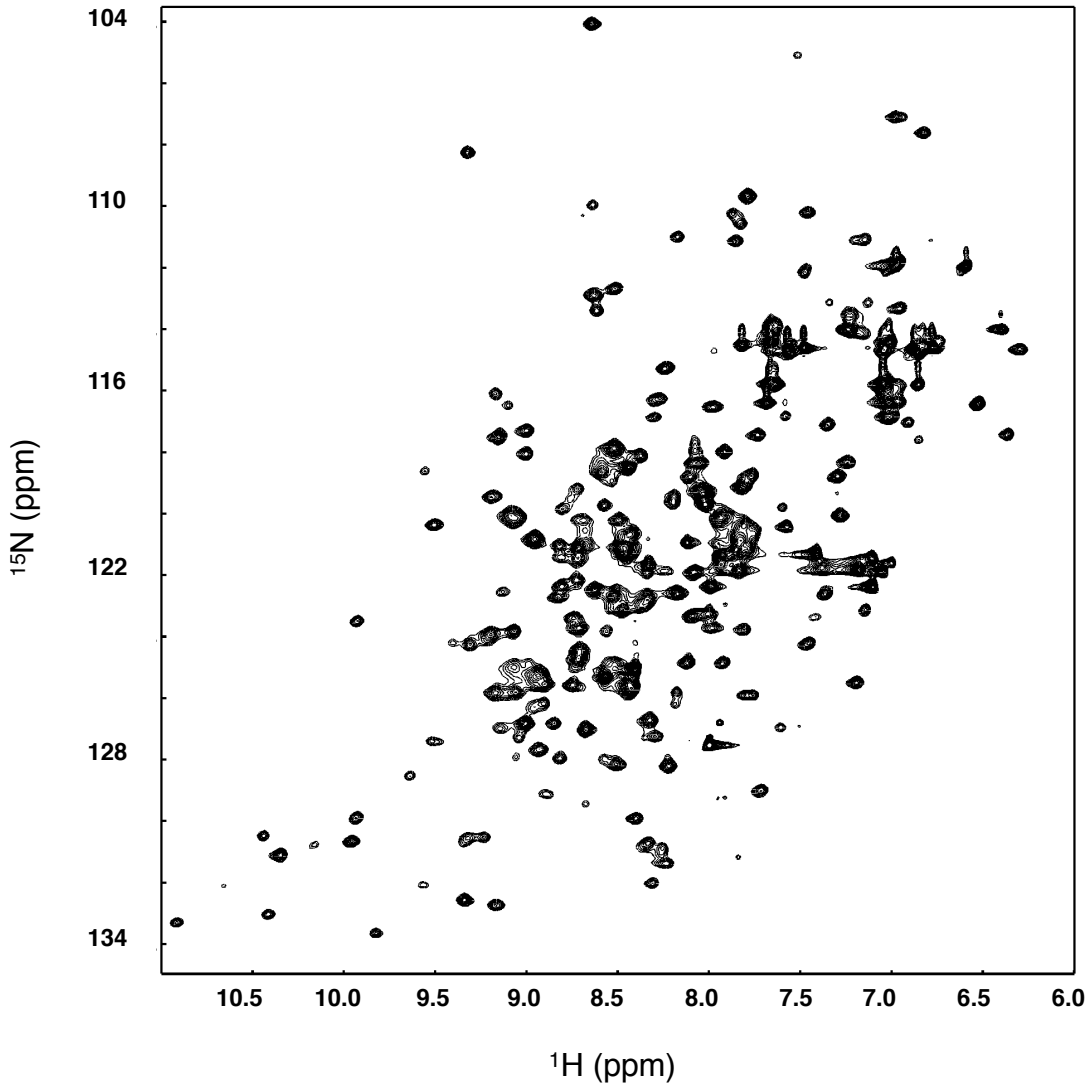



**Figure 2.5: $^{1}H^{15}N$- HSQC spectra of HPX-1.** The spectrum was recorded in 20 mM NaOAc pH 4.8, 5 mM $CaCl_2$, 50 µM $ZnCl_2$, 0.02% (w/v) $NaN_3$ and 10% $D_2O$ at 25°C and at a concentration of 283 µM on an Agilent Inova 600 MHz spectrophotometer.

## 2.4.2 proCAT-3 protein

The pET-28b/proCAT-3 plasmid was used to express C-terminally truncated human proCAT-3 as an N-terminal hexahistidine ($H_6$) tagged fusion protein, thereby allowing purification of the target protein by immobilised metal affinity chromatography (IMAC). Small-scale protein expression was trialled in the *E. coli* BL21-CodonPlus (DE3)-RIPL strain (Section 2.3.7) and analysis of pre- and post-induction expression indicated the presence of a protein (~ 33 kDa) corresponding to $H_6$-proCAT-3. Further analysis indicated that $H_6$-proCAT-3 was expressed in approximately equal amounts in both the soluble and insoluble fractions. Despite establishment of a suitable refolding protocol to

recover and reconstitute proCAT-3 by Dr Stephen Prior (detailed in Section 2.3.13), the assays in this study utilise only purified protein derived from the soluble fraction as shown in Figure 2.6. Therefore, protein-expressing cells were grown overnight at 20 °C to maximise protein yields during large-scale $H_6$-proCAT-3 preparations. Subsequent IMAC was used to purify the $H_6$-tagged target protein from the clarified crude cell extract ensuring significant reduction of contaminating proteins.

Activation of $H_6$-proCAT-3 was initially achieved by incubation with chymotrypsin (1:60 molar concentration, respectively) for 10 minutes at 37 °C (data not shown). However, this method was swiftly abandoned and replaced with the APMA activation protocol (Section 2.3.21), which removed the need to reintroduce an active protease, instead using autolysis to ensure PRO domain removal and convenient N-terminal $H_6$-tag removal. Buffer-exchanged enzymatic digestion reactions were desalted and separation of post-digest products was achieved using size exclusion chromatography. This purification strategy yielded ~ 2 mg of >99% homogeneous active CAT-3 per 500 mL of LB media (Table 2.9).

Activity of purified CAT-3 using a fluorogenic peptide substrate assay (Section 2.3.24) was taken as suitable evidence for use in proMMP-1* activation (Figure 2.7).

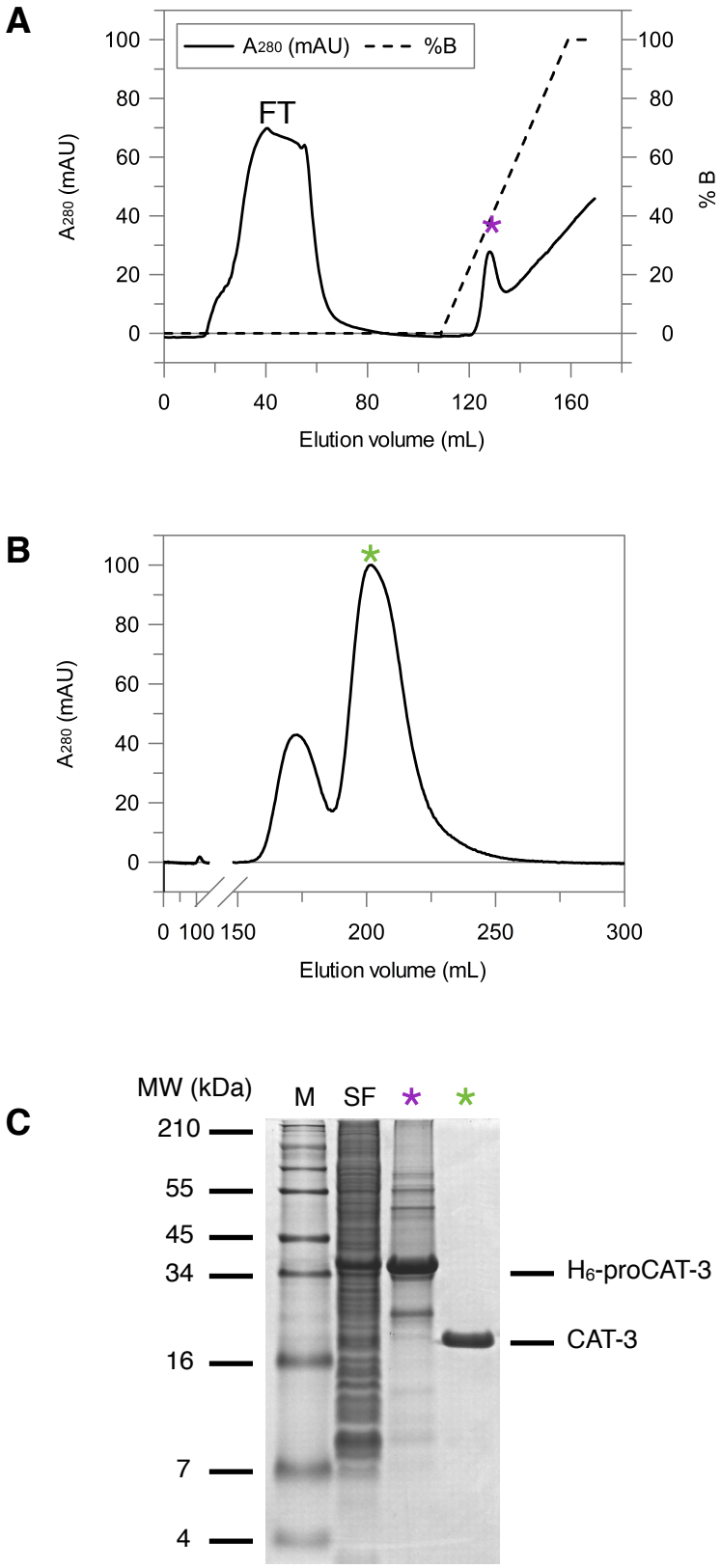



**Figure 2.6: proCAT-3 purification and activation.** $H_6$-proCAT-3 expressed in the soluble fraction (**SF**) was applied to an Ni Sepharose HisTrap HP column at a flow rate of 5 mL/min (**A**). Bound protein was eluted with a 0-500 mM gradient of imidazole. Pooled fractions spanning the peak (*) were then incubated with 0.5 mM APMA, desalted and applied at 1 mL/min to a pre-equilibrated HiLoad Superdex 75 26/60 prep grade column (**B**). Fractions containing activated CAT-3 (*) were visualised alongside other stages of the purification strategy on a 16.5% (w/v) Tris-Tricine SDS-PAGE gel (**C**). Molecular weight was estimated using SeeBlue Plus2 prestained molecular weight marker (**M**).

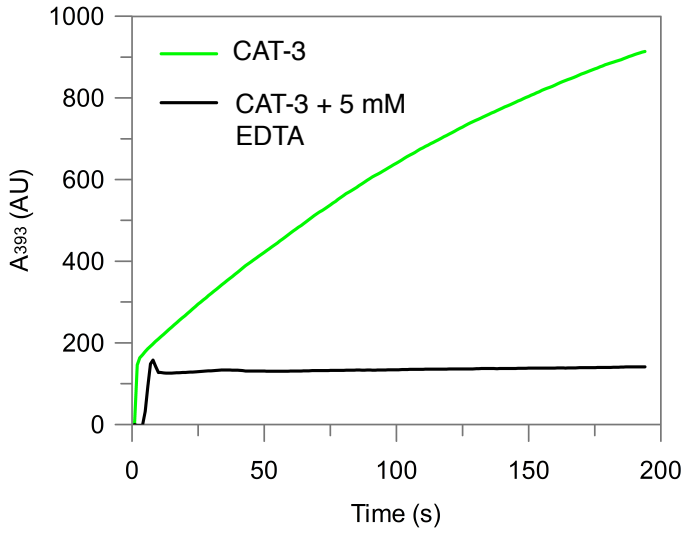



**Figure 2.7: CAT-3 activity against fluorogenic peptide substrate (Mca-Arg-Pro-Lys-Pro-Val-Glu-Nva-Trp-Arg-Lys-(Dnp)-$NH_2$).** Assays were performed at $\lambda_{ex}$ 325 nm and $\lambda_{em}$ 393 nm. Detected changes in fluorescence is in arbitrary units (**AU**). Typical activity observed at a concentration of 10 nM is shown in *green*. Lack of metalloproteinase activity was observed with EDTA treatment (*black*).

## 2.4.3 proMMP-1 protein

### 2.4.3.1 Hydrolytically impaired proMMP-1*

The pET-3a/proMMP-1* construct, containing the sequence coding for the full-length human proMMP-1 protein was used for transformations in both *E. coli* DH5α and BL21-CodonPlus (DE3)-RIPL strains.

High level protein expression was induced in the BL21-CodonPlus (DE3)-RIPL expression strain with 0.4 mM IPTG as visualised by SDS-PAGE analysis (Figure 2.8A). Despite the strong selectivity of the T7 promoter for its phage-encoded polymerase, residual basal level expression from this pET-3a construct was clearly present in the absence of IPTG treatment. Furthermore, deposition of full-length protein in bacterial inclusion bodies (Figure 2.8B) appeared to protect the cells from any potentially toxic effects. The inclusion body preparation protocol used during HPX-1 production (Section 2.3.12) was adopted, and again high concentrations of full-length protein was isolated in this manner, although it was clear that higher background protein levels were concomitant with overnight protein expression at 20 °C. However, accumulation of precipitated bacterial proteins during the refolding procedure, which were removed by filtration, appeared to remove many of these contaminants as well as misfolded proMMP-1 aggregates. Refolded protein samples were further purified by cation

exchange chromatography (Figure 2.9); this step also served to concentrate large dilute refolding volumes.

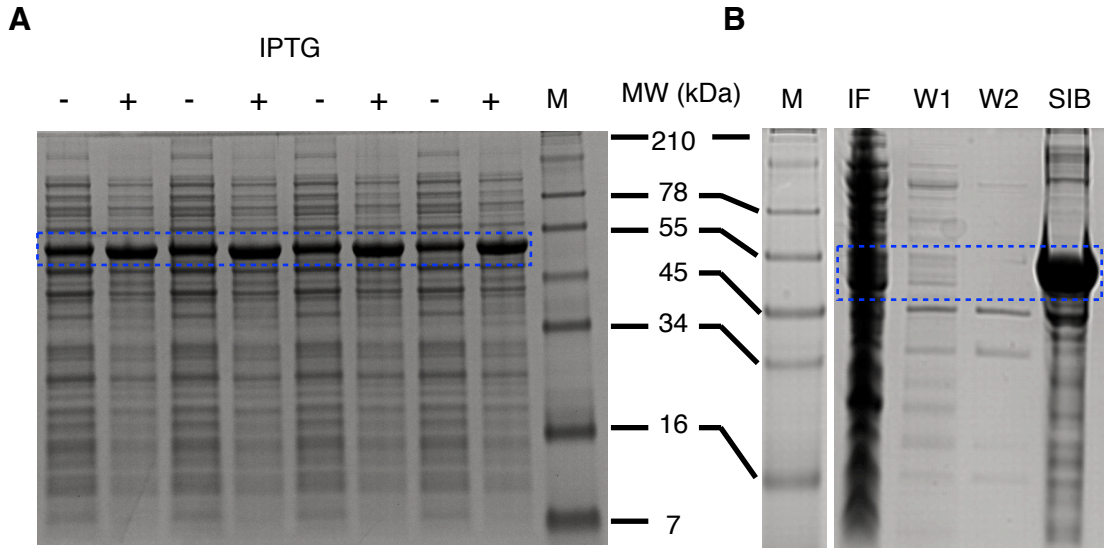



**Figure 2.8: SDS-PAGE analysis of proMMP-1* expression and Inclusion Body preparation.**
(**A**) Expression of a protein of ~ 53 kDa (HPX-1) was induced upon treatment with 0.4 mM IPTG and visualised on a 16.5% (w/v) Tris-Tricine reducing gel, indicated by the *dashed blue box*. (**B**) Insoluble material was harvested by centrifugation (**IF**) and washed (**W1** and **W2**). Pelleted inclusion bodies were then solubilised overnight (**SIB**). All samples were ran alongside SeeBlue Plus2 prestained molecular weight marker (**M**).

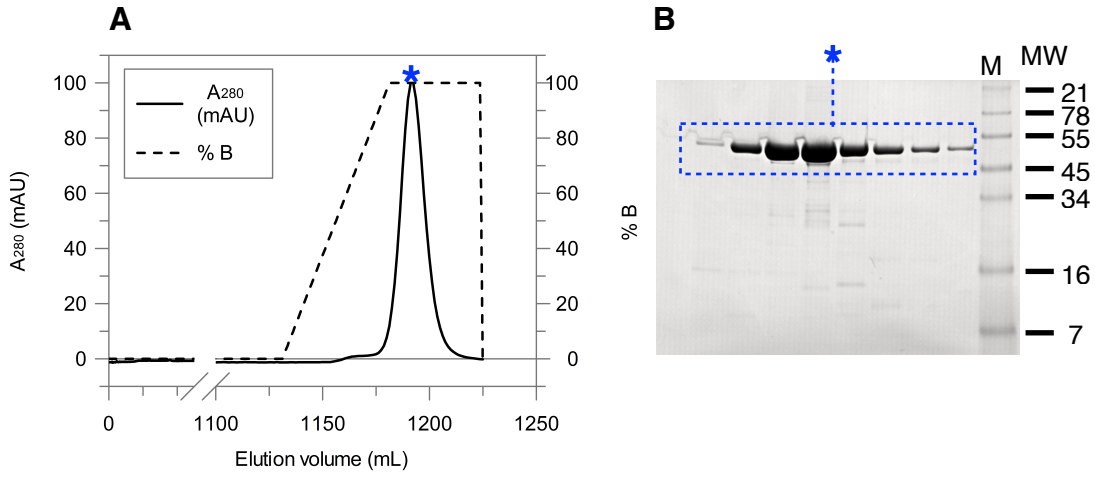



**Figure 2.9: Cation exchange purification of proMMP-1*.**
(**A**) Refolded protein solutions were applied to a pre-equilibrated 30 mL SP-Sepharose FF XK26 column at a flow rate of 10 mL/min. Bound protein was eluted with a 0-1 M NaCl gradient. (**B**) Pooled fractions spanning the peak (*) were visualised on a 16.5% (w/v) Tris-Tricine SDS-PAGE gel. Molecular weight was estimated using SeeBlue Plus2 prestained molecular weight marker (**M**).

Final purification of proMMP-1* was achieved using gel filtration chromatography (Figure 2.10A) and resulted in samples of >99% homogeneity. The final purification strategy yielded ~ 6 mg of proMMP-1* (Table 2.9).

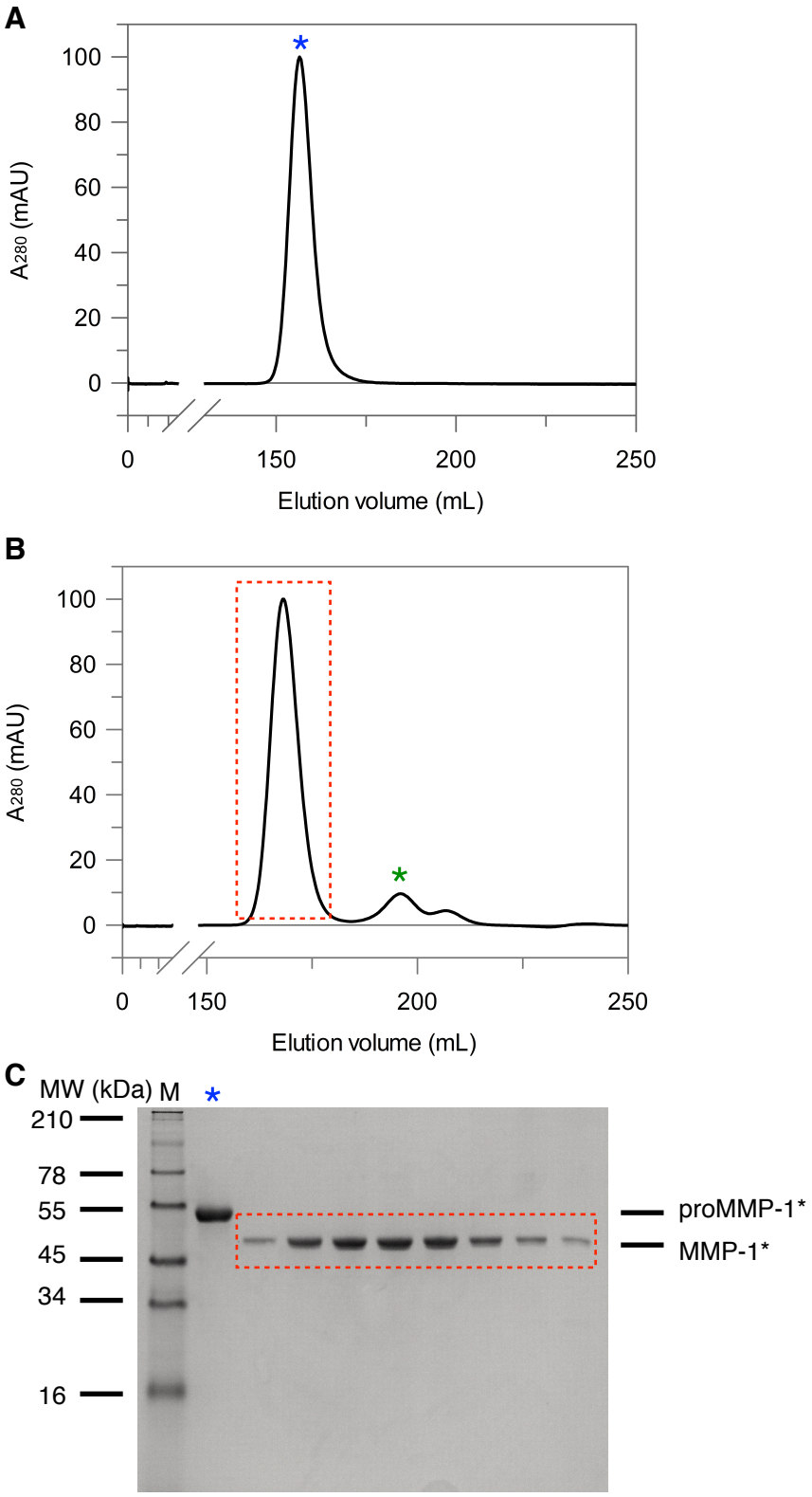



**Figure 2.10: Gel filtration chromatography of proMMP-1* and activated MMP-1*.** (**A**) CEX purified proMMP-1* was applied to a pre-equilibrated HiLoad Superdex 75 16/60 prep grade column column at a flow rate of 1 mL/min. Purified proMMP-1* is indicated with a *blue asterisk* (*). (**B**) Purified proMMP-1* was incubated with 0.1 mM APMA and CAT-3 (*) at a 10:1 molar concentration, desalted and purified by gel filtration chromatography. Purified MMP-1* is shown with a *dashed red box*. (**C**) proMMP-1* and MMP-1* samples purified by gel filtration chromatography were visualised on a 16.5% (w/v) Tris-Tricine SDS-PAGE reducing gel. Molecular weight was estimated using SeeBlue Plus2 prestained molecular weight marker (**M**).

### 2.4.3.1.1 Maturation of proMMP-1*

Destabilisation and cleavage of the PRO-domain following enzymatic processing has been well-reported (Nagase, Enghild et al. 1990, Suzuki, Enghild et al. 1990). Initially, small-scale MMP-1 maturation was attempted using enzymatic digestion with trypsin, which led to molecular weight reduction and production of a 46 kDa intermediate (Figure 2.11A). Impeded by inherent lack of autolytic activity in MMP-1*, and despite repeated attempts to optimize this activation method by altering both incubation length and trypsin concentration, the final yield of mature MMP-1* was insufficient for further studies. As an alternative, digestion with CAT-3 instead indicated specific cleavage of the final Gln80 and Phe81 bond yielding a ~42 kDa product, suggestive of PRO-domain loss and hence maturation of the enzyme (Figure 2.11B). However, incubation length studies of proMMP-1* in the presence of CAT-3 at equimolar concentrations, did indicate a requirement for extensive incubation times (>24 hours) and was further hampered during gel filtration chromatography, which was unable to separate post-digestion products. Contrastly, it was observed that treatment with both APMA (Section 2.3.21) and CAT-3 allowed a reduction in both incubation length (4 hours) and in molar ratio of the activating protease (1:10). The result of purification of mature MMP-1* by gel filtration chromatography is shown in Figure 2.10B and 2.10C.

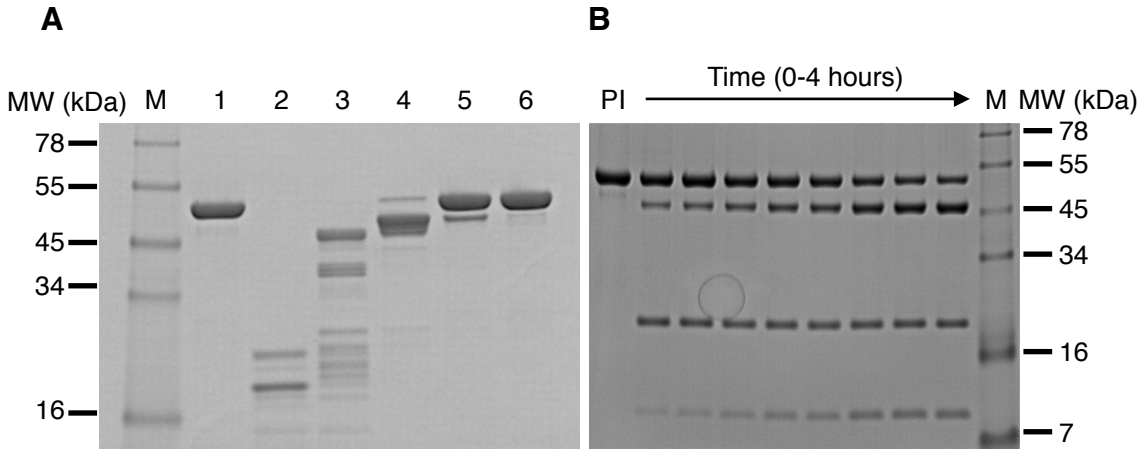



**Figure 2.11: Protease activation of proMMP-1*.** Cleavage by trypsin yields an intermediate 46 kDa product. The final cleavage between Gln80 and Phe81 by CAT-3 yields mature MMP-1 (42 kDa) (Murphy, Cockett et al. 1987). (**A**) Enzymatic digestion of proMMP-1* using serially diluted trypsin visualised on a 12% (w/v) Tris-Tricine SDS-PAGE gel. Molecular weight was estimated using SeeBlue Plus2 prestained protein ladder (**M**). Pre-digested proMMP-1* control (*Lane 1*), proMMP-1* digested with trypsin-270 mg/mL (*Lane 2*), proMMP-1* digested with trypsin 27 mg/mL (*Lane 3*), proMMP-1* digested with trypsin-2.7 mg/mL (*Lane 4*), proMMP-1* digested with trypsin-270 ng/mL (*Lane 5*), proMMP-1* digested with trypsin-27 ng/mL (*Lane 6*). (**B**) Enzymatic digestion timecourse (0-4 hours) of proMMP-1* using CAT-3 visualised on a 12% (w/v) Tris-Tricine SDS-PAGE gel. An untreated proMMP-1* sample is shown for comparison (**PI**).

### 2.4.3.2 Hydrolytically active MMP-1 proteins

#### 2.4.3.2.1 CAT-1

Fortuitously, it was found that active CAT-1 could be generated by exploiting the autolytic activity of MMP-1 and purified simply with IMAC chromatography. Overnight incubation at 37°C was found sufficient to observe the presence of all major MMP-1 components due to autolysis. The mass similarity between the CAT and HPX domain precluded subsequent purification by gel filtration chromatography. However, application to a column prepacked with Ni Sepharose 6 Fast Flow resin separated the major domain components (Figure 2.12). Unexpectedly, the CAT domain was found in the flowthrough while the HPX domain was retained on-column requiring an imidazole wash for removal. This purification method was found to consistently separate the CAT and HPX domains, allowing CAT-1 isolation for hydrolytic and limited proteolysis studies.

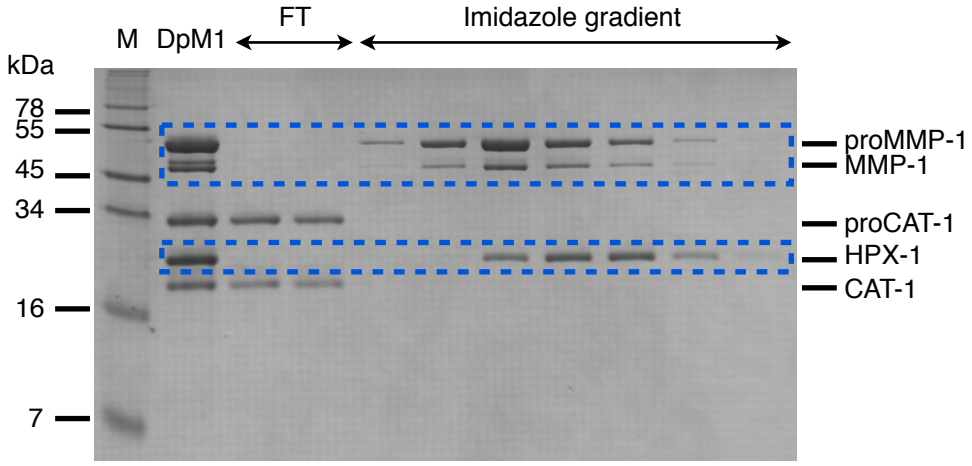



**Figure 2.12: Nickel binding by HPX-1.** Degraded proMMP-1 (**DpM1**) is separated by IMAC chromatography. All HPX-1 containing fragments (*blue dashed box*) are retained on column. The absence of HPX-1 prevents binding to a nickel-containing resin (**FT**). Samples were visualised on a 12% (w/v) Tris-Tricine SDS-PAGE gel. Molecular weight was estimated using SeeBlue Plus2 prestained protein ladder (**M**).

### 2.4.3.2.2 proMMP-1

Generation of the proMMP-1 revertant by mutating E219A back to glutamate was completed using SDM of the proMMP-1* pET-3a plasmid using the Quikchange® II XL Site-Directed Mutagenesis Kit (Agilent Technologies). The mutagenic efficiency level was 88%. Clones potentially harbouring mutants plasmids were randomly selected and harvested using a QIAprep Spin Miniprep Kit prior to dilution to 100 ng/μL in preparation for sequencing. DNA sequence analysis confirmed successful incorporation of mutagenic bases and hence substitution of the alanine residue with the glutamate residue required for catalytic activity in MMP-1 (See Appendix).

Hydrolytically active proMMP-1 was produced in an identical fashion to that of proMMP-1*, and subsequently yielded similar levels of purified protein. The only notable observation was an increased level of proteolysis, with basal levels of enzyme maturation occurring when stored at concentrations exceeding 0.5 mg/mL. Storage at 4 $^{o}$C reduced but did not prevent proteolysis.

Treatment with APMA (Section 2.3.21) was used to maximise the final yield of MMP-1 prior to purification by gel filtration chromatography (Section 2.3.22). Activity of mature MMP-1 against a chromogenic MMP substrate (Section 2.3.25) was compared with that of the MMP-1 catalytic domain to ensure the refolding protocol had not adversely impacted catalytic activity (Figure 2.13A). Activity against the complex substrate Type I collagen was also assessed (Figure 2.13B).

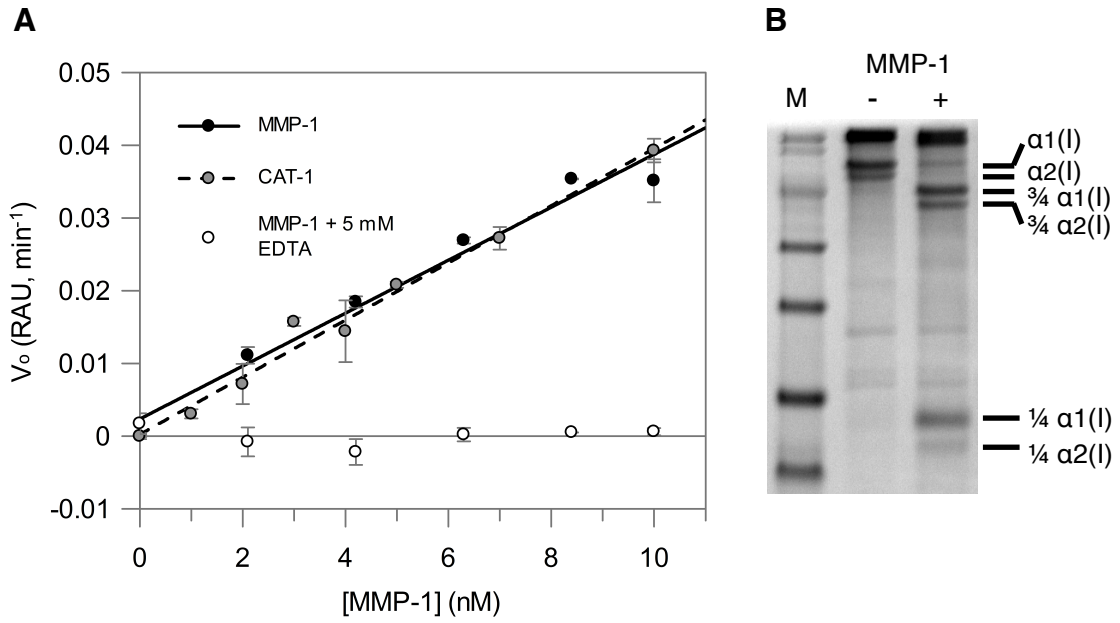



**Figure 2.13: MMP-1 activity.** (**A**) MMP-1 (wild-type) activity against chromogenic peptide substrate (Ac-Pro-Leu-Gly-[2-mercapto-4-methyl-pentanoyl]-Leu-Gly-$OC_2H_5$). Assays were monitored at λ 412 nm at 37 °C. Lack of metalloproteinase activity was observed with 5 mM EDTA treatment. (**B**) Type I collagen hydrolysis. Collagen samples were incubated overnight at 37 °C in the presence and absence of MMP-1 (1:30 molar ratio of MMP-1) and visualised on a 12% (w/v) Tris-Tricine SDS-PAGE alongside SeeBlue Plus2 prestained molecular weight marker (**M**) that had been run at 150 V for 150 minutes.

### 2.4.4 proCAT-1* protein

Two constructs containing C-terminally truncated inactive human proMMP-1 were initially tested for inducible protein expression. In expression trials (Section 2.3.7), the untagged protein (pET-3a/proCAT-1* construct) was poorly expressed and showed no increase in expression levels upon induction with 0.4 mM IPTG. In contrast, the 29.4 kDa N-terminal hexahistidine tagged fusion protein (pET-28b/$H_6$-proCAT-1* construct) showed inducible expression of soluble protein (Figure 2.14A). Therefore, the residual basal expression, observed previously for proMMP-1 (Figure 2.8A) using the pET-3a plasmid system, may well have detrimental toxic effects.

The N-terminal hexahistidine tag was successfully removed using thrombin (Section 2.3.16). Clarified lysate was first applied to a HisTrap FF column, then 25 U thrombin was manually injected onto the column and incubated overnight at 37 °C. After the digest, 5.4 mL p-amidobenzamidine column was attached downstream of the HisTrap FF column and both columns were then washed with 5 column volumes of buffer to elute proCAT-1* whilst trapping the thrombin. The uncut $H_6$-proCAT-1* was eluted alongside other bacterial contaminants with a 0-500 mM imidazole gradient

(Figure 2.13B). Despite significant removal of unbound material prior to proteolysis, bacterial contaminants continued to persist, most likely generated by non-specific thrombin cleavage. Anion exchange chromatography was used to successfully remove virtually all contaminants (Figure 2.14C) and subsequent use of gel filtration chromatography resulted in samples of >99% homogeneity (Figure 2.15) with a yield of ~ 5 mg of protein per 500 mL of LB media (Table 2.9).

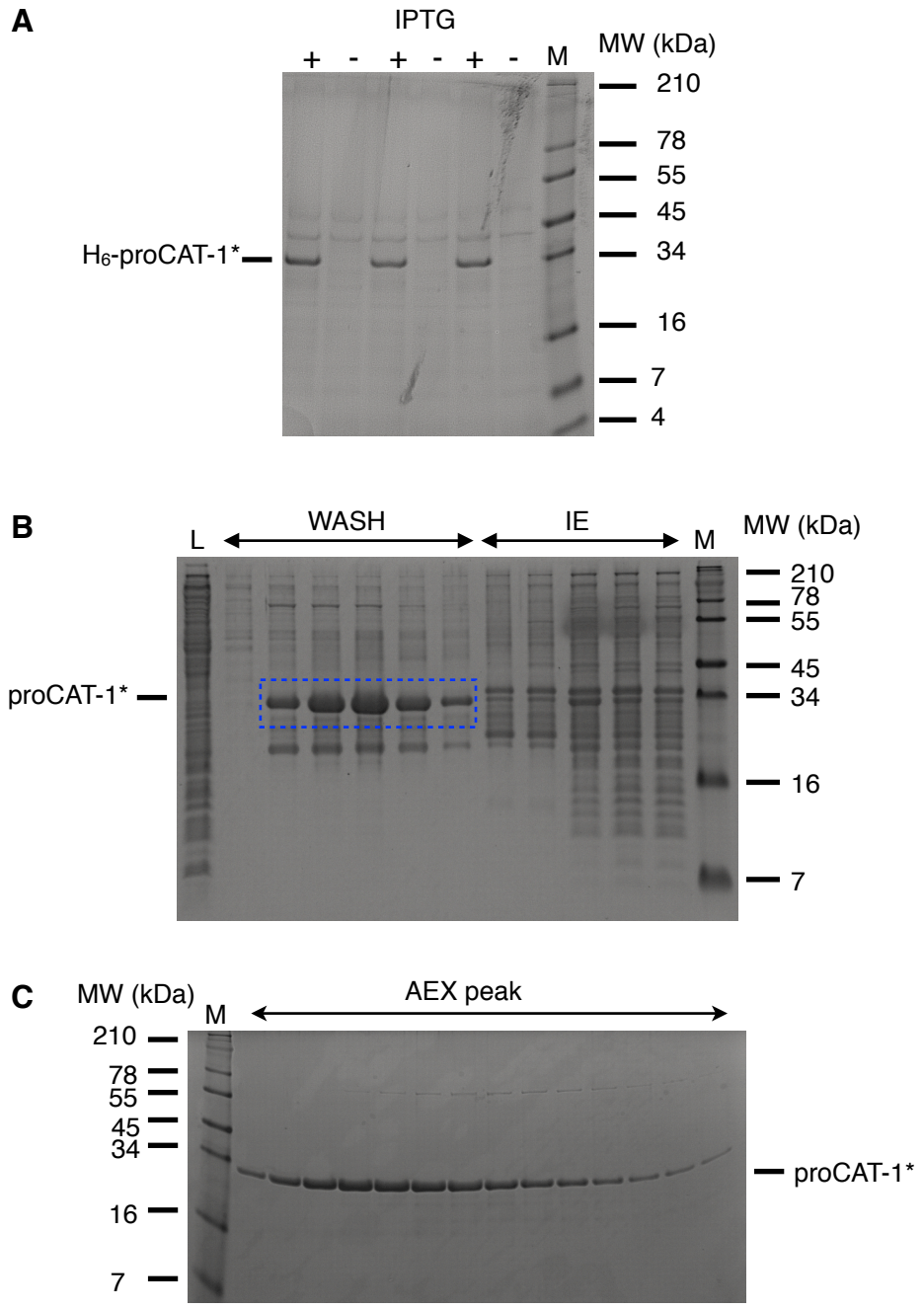



**Figure 2.14: SDS-PAGE analysis of $H_6$-proCAT-1* expression, tag-removal and purification.** (**A**) Expression of a protein of ~ 30 kDa ($H_6$-proCAT-1*) was induced upon treatment with 0.4 mM IPTG and visualised on a 16.5% (w/v) Tris-Tricine SDS-PAGE reducing gel. (**B**) Cell lysate (**L**) was incubated with thrombin overnight at 37 °C. Cleaved proCAT-1* (*dashed blue box*) was eluted with 5 CV of buffer (**WASH**). Uncleaved bound material was eluted with a 0-500 mM imidazole gradient (**IE**). (**C**) Samples were then applied to a pre-equilibrated Q-Sepharose XK26 column at 1 mL/min. Bound protein was eluted with a 0-1 M NaCl gradient (**AEX peak**). All samples were ran alongside SeeBlue Plus2 prestained molecular weight marker (**M**).

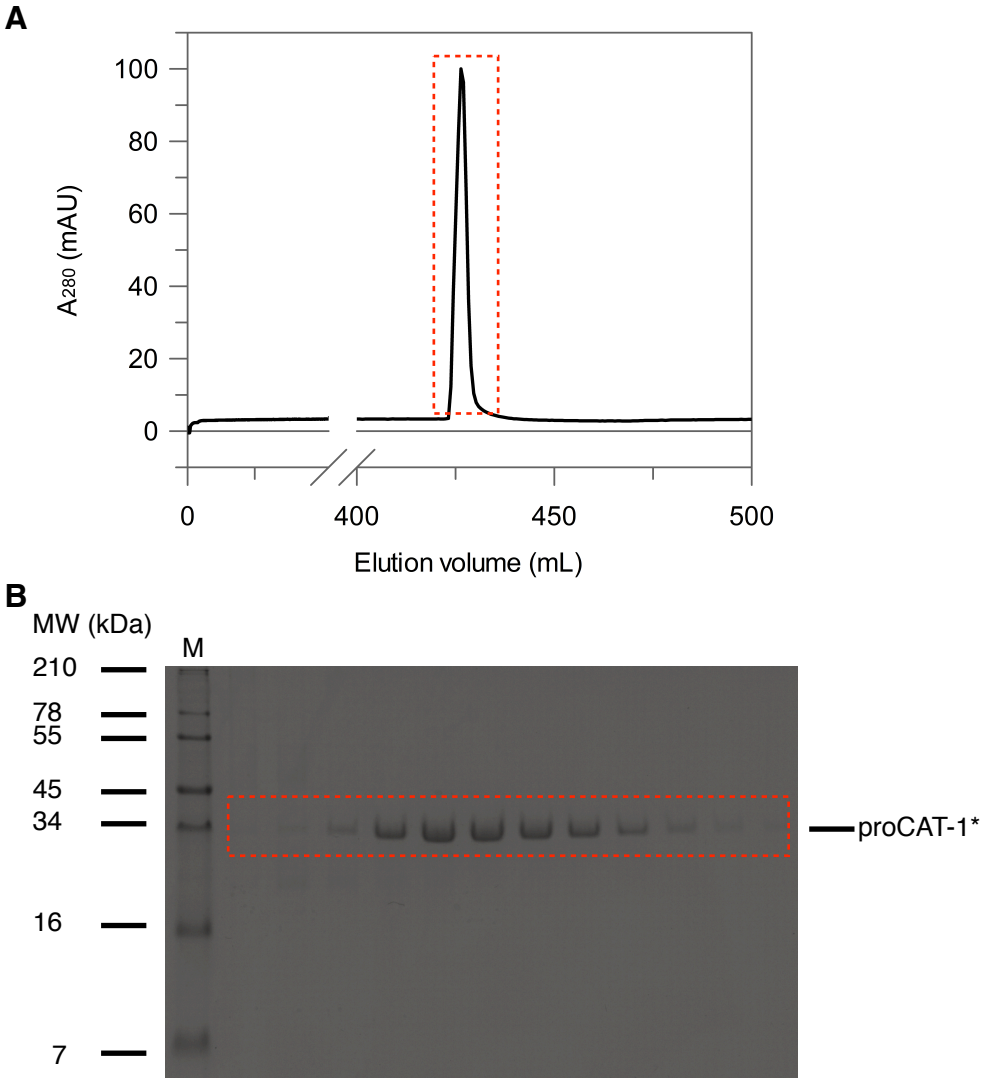



**Figure 2.15: Gel filtration chromatography of proCAT-1*.** (**A**) AEX purified proCAT-1* was applied to a pre-equilibrated HiLoad Superdex 75 16/60 prep grade column column at a flow rate of 1 mL/min. The eluted peak was visualised on a 16.5% (w/v) SDS-PAGE gel (**B**) The eluted peak containing purified proCAT-1* (*dashed red box*) was visualised on a 16.5% (w/v) SDS-PAGE gel. Molecular weight was estimated using SeeBlue Plus2 prestained molecular weight marker (**M**).

The conformation of proCAT-1* was assessed by acquisition of a 2D $^{1}$H-$^{15}$N HSQC spectrum (Section 2.3.29). Like the HPX-1 $^{1}$H-$^{15}$N HSQC spectrum shown in Section 2.4.1, proCAT-1* also exhibited chemical shift dispersion characteristic of a folded protein (Figure 2.16A).

The exact role of APMA during the activation process remains unknown. Therefore, a $^{1}$H-$^{15}$N HSQC spectrum was acquired to ensure that no detrimental structural effects occur after APMA treatment. The data showed that proCAT-1* remained folded with a number of changes in crosspeak position (Figure 2.16B), which suggests a mechanism of interaction that involves perturbation of the PRO domain.

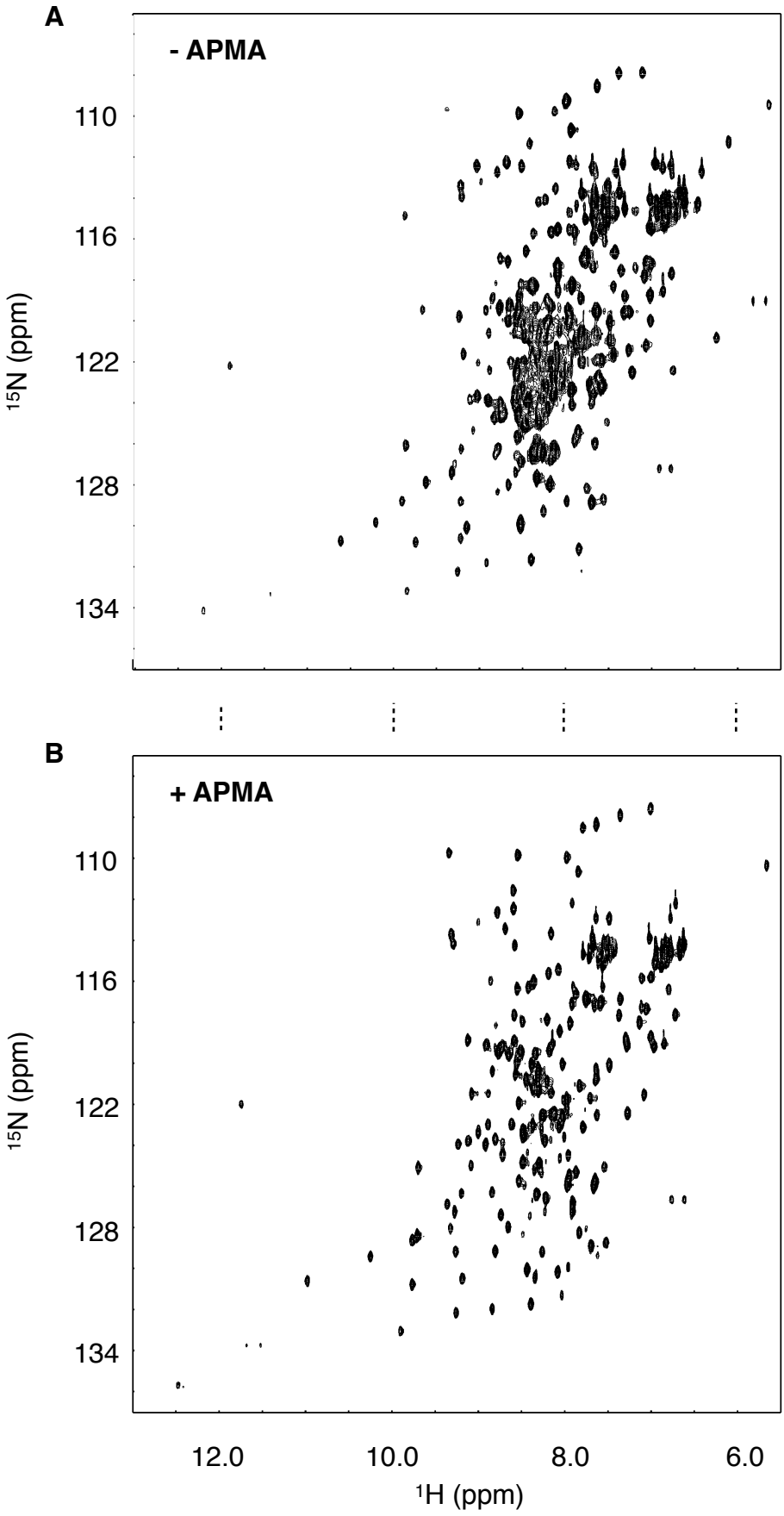



**Figure 2.16: $^{1}H^{15}N$- HSQC spectra of proCAT-1*.** (A) The spectrum of proCAT-1* (916 µM) prior to APMA treatment. (B) The spectrum of proCAT-1* (786 µM) following a 4 hour treatment with 0.5 mM APMA. Residual APMA was removed using a HiPrep desalting column prior to spectral acquisition. All spectra were recorded in 10 mM Tris-HCl pH 7.2, 5 mM $CaCl_2$, 150 mM NaCl, 0.02% (w/v) $NaN_3$ and10% $D_2O$ pH 7.4 at 25°C on a Agilent Inova 600 MHz spectrophotometer.

## 2.5 Conclusions

The protocols developed and described herein were found to be both robust and reproducible, producing milligram quantities of HPX-1, proMMP-1, CAT-3 and catalytically-impaired proCAT-1* and proMMP-1*; suitable for all structural and functional assays described in later chapters. Production of full-length proMMP-1* proved most challenging. Initially, an N-terminal hexahistidine tagged fusion protein (pET-3a/$H_6$-proMMP-1* construct) was expressed but consistently precipitated during varying refolding protocols. In contrast, the tag-free proMMP-1* protein was refolded and purified successfully.

Activation of proMMP-1 by MMP-3 is well described (Murphy, Cockett et al. 1987). In this study, proCAT-3 was expressed, purified and activated for this purpose. APMA-initiated autolysis resulted in acquisition of full MMP enzymatic activity. Consequently, CAT-3 had catalytic activity against a fluorogenic peptide, and when incubated with proMMP-1* appeared to cleave at a specific bond, most likely the Q99-F100 bond (Van Wart and Birkedal-Hansen 1990). Limited proteolysis with increasing concentrations of trypsin indicated the presence of MMP-1* activation intermediates, which is consistent with previous studies (Nagase, Enghild et al. 1990, Suzuki, Enghild et al. 1990). Resistance to non-specific protease digestion by both CAT-3 and trypsin suggests that proMMP-1* possesses significant tertiary structure.

The protein mass constraints of heteronuclear NMR spectroscopy unfortunately prohibited analysis of full-length proMMP-1 at the current University of Portsmouth NMR facility, however analysis of both proCAT-1* and HPX-1 domains in isolation indicated possession of folded tertiary structure in all three domains and concurs with the protease resistance data. In the case of HPX-1, which requires refolding, the NMR data is particularly important when assessing the impact of mutated amino acid residues on substrate binding (Chapter 3). Interestingly, NMR analysis indicated structural perturbation of proCAT-1* when treated with APMA. This observation supports an activation model that requires domain destabilisation rather than residue modification within the PRO domain alone. Furthermore, the NMR data shown here supports previous studies of proMMP-2 and -3 activation (Chen, Noelken et al. 1993, Itoh, Binner et al. 1995, Galazka, Windsor et al. 1996), which appear to require zymogen perturbation.

Restoration of full MMP-1 enzymatic activity was achieved using site-directed mutagenesis to replace the necessary glutamate residue. In the presence of APMA, this

revertant was able to remove the inhibitory PRO domain by autolysis and displayed activity against a simple chromogenic peptide. More importantly, refolded MMP-1 exhibited activity against type I collagen, the complex natural substrate. This suggests that all necessary determinants, both CAT-1 and HPX-1 (Chung, Yoshida et al. 2004), are present and functional.

# CHAPTER 3: IDENTIFYING THE COLLAGEN-BINDING SITE OF THE MMP-1 HEMOPEXIN DOMAIN BY MUTAGENESIS AND SURFACE PLASMON RESONANCE

## 3.1 Introduction

Fibrillar collagen degradation by the prototypic collagenase, MMP-1, results from the cooperative action of both the CAT and HPX domains (Chung, Yoshida et al. 2004). Previous studies have implicated the accessory HPX domain in binding the collagen substrate in preparation for hydrolysis of the Gly775-Ile776 or Gly775-Leu776 scissile bond by the CAT domain. Despite the lack of obvious binding clefts or pockets on the HPX surface (Li, Brick et al. 1995, Jozic, Bourenkov et al. 2005, Iyer, Visse et al. 2006), the existence of 'exosites' necessary for substrate recognition and cleavage have been predicted (Overall 2002). A recent study examining the interaction between an MMP-1 active site mutant (E219A) and a synthetic triple helical peptide (THP) using hydrogen/deuterium exchange mass spectrometry (HD/MS) implicated blades 1 and 4 of the HPX domain in collagen binding (Lauer-Fields, Chalmers et al. 2009). Subsequent mutagenesis and assay identified residues Ile290 and Arg291 in the A-B loop of HPX-1 blade 1 as an exosite for collagenolysis. However, this 'exosite' appears to be much more extensive. An NMR-monitored ligand titration of $^{15}$N-labeled HPX-1 with unlabeled α1(I)772-787 THP conducted in the laboratory of Dr Andrew Pickford, showed significant line-broadening indicative of substrate binding for residues Arg291-Glu293, Phe308-Glu311, Glu313-Phe316, Ser318-Val319 and Phe320-Gly328 of blade 1, and Gly353-Val356 of blade 2 (Figure 3.1).

This study uses a comprehensive programme of alanine-scanning site-directed mutagenesis and assay to study the complex formed between the HPX-1 domain and a synthetic THP that encompasses the MMP-1 cleavage site of the collagen α1(I) chain. Surface Plasmon Resonance (SPR) was used to examine the functional contribution of 12 individual amino acid residues to the putative collagen binding site located in blades 1 and 2 of HPX-1, while heteronuclear NMR spectroscopy was used to investigate selected HPX-1 point mutants to ensure that native structure was retained and unimpaired by mutagenesis.

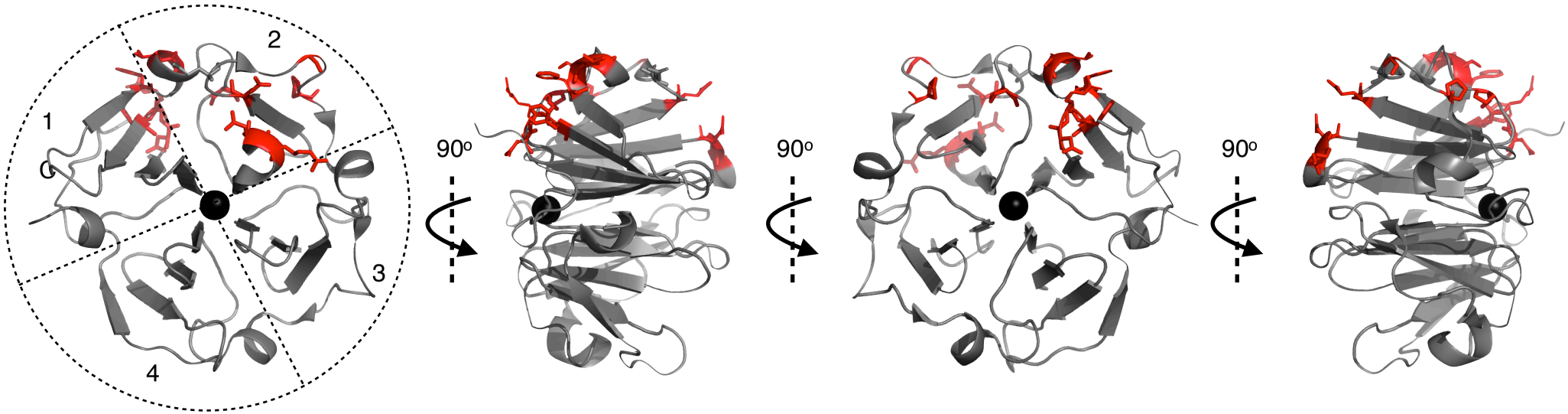



**Figure 3.1: Ribbon representation of the crystal structure of HPX-1.** Putative 'exosite' residues identified using an NMR-monitored ligand titration of $^{15}$N-labelled HPX-1 with unlabeled α1(I)772-787 THP are coloured in *red*. The four blades of the β-propeller are outlined and labeled *1– 4*.

Table 3.1: SDM candidate residues[a]

| Residue | Selection justification | Residue | Selection justification | Residue | Selection justification | Residue | Selection justification |
|---|---|---|---|---|---|---|---|
| D299 | 17.6% reduction in peak volume upon THP addition | F316 | 20.9% reduction in peak volume upon THP addition | R337 | 30.1% reduction in peak volume upon THP addition | Q354 | Adjacent to highly perturbed residue G353 (34.1%). |
| R300 | Proximity to D299 | S318 | Positioned on α-turn and in proximity to F316 and V319 | D338 | Proximity to R337 | H358 | Proximity to Q352 and Q354 |
| F301 | Exposed hydrophobic residue in proximity to D299 | V319 | 20.4% reduction in peak volume upon THP addition | Q352 | Adjacent to highly perturbed residue G353 (34.1%). | P361 | Exposed position and proximity to perturbed residue P325 |

[a] data taken from Arnold, L.H. (2010). *Biophysical Characterization of Collagen Binding by the Hemopexin Domain of Matrix Metalloproteinase-1 (MMP-1).* Ph.D. Thesis. University of Portsmouth: UK

### 3.1.1 The Surface Plasmon Resonance phenomenon

An evanescent wave is generated when the total internal reflection of p-polarized light (polarization occurs parallel to the plane of incidence) occurs at the interface between a medium of high refractive index and a medium of low refractive index. The amplitude of the evanescent wave decays exponentially away from the interface, but can be enhanced using a suitable-conducting thin film of inert metal at the interface, typically gold. Optical excitation at a certain incident angle transfers energy to surface plasmons, collective electron oscillations; enhancing the evanescent electric field amplitude, penetrating the low refractive index medium, and thereby reducing the reflected light intensity. Consequently, the incident angle (SPR angle) at which plasmon excitation (resonance) occurs is extremely sensitive to any change in the refractive index of the medium adjacent to the metal surface. Mass bound to the surface alters the refractive index and therefore the SPR angle. Thus, SPR is an optical phenomenon that can be exploited to observe interactions between immobilized biomolecules and their cognate binding partners.

### 3.1.2 Biacore Technology

SPR-based biosensor assays offer rapid screening of molecular interactions in real time. Instruments have three primary components: an SPR optical detection system, a microfluidic flow cell system and a gold-dextran sensor surface, which together enable detection and quantitative measurements of binding interactions in real time (Components of Biacore systems are fully described in the Biacore Sensor Surface Handbook - GE Healthcare, 2008). A prism coupling p-polarized light, the gold film and a light-detecting device in the Kretschmann configuration (Figure 3.2) is the most widely used optical system. Biacore instruments utilise this arrangement to measure variations in the refractive index. The subsequent shift of the SPR angle is proportional to the mass change at the sensor surface determined by the binding and dissociation of biomolecules. This is monitored using the Biacore control software and is visualised as a sensorgram (Figure 3.2) with response units (RU), corresponding to the angle shift, plotted as a function of time. Highly sensitive instruments like the Biacore T-200 can reliably detect changes of 1 RU, which corresponds to approximately 1 pg/mm$^2$ of mass on the sensor surface area (GE Healthcare, 2008).

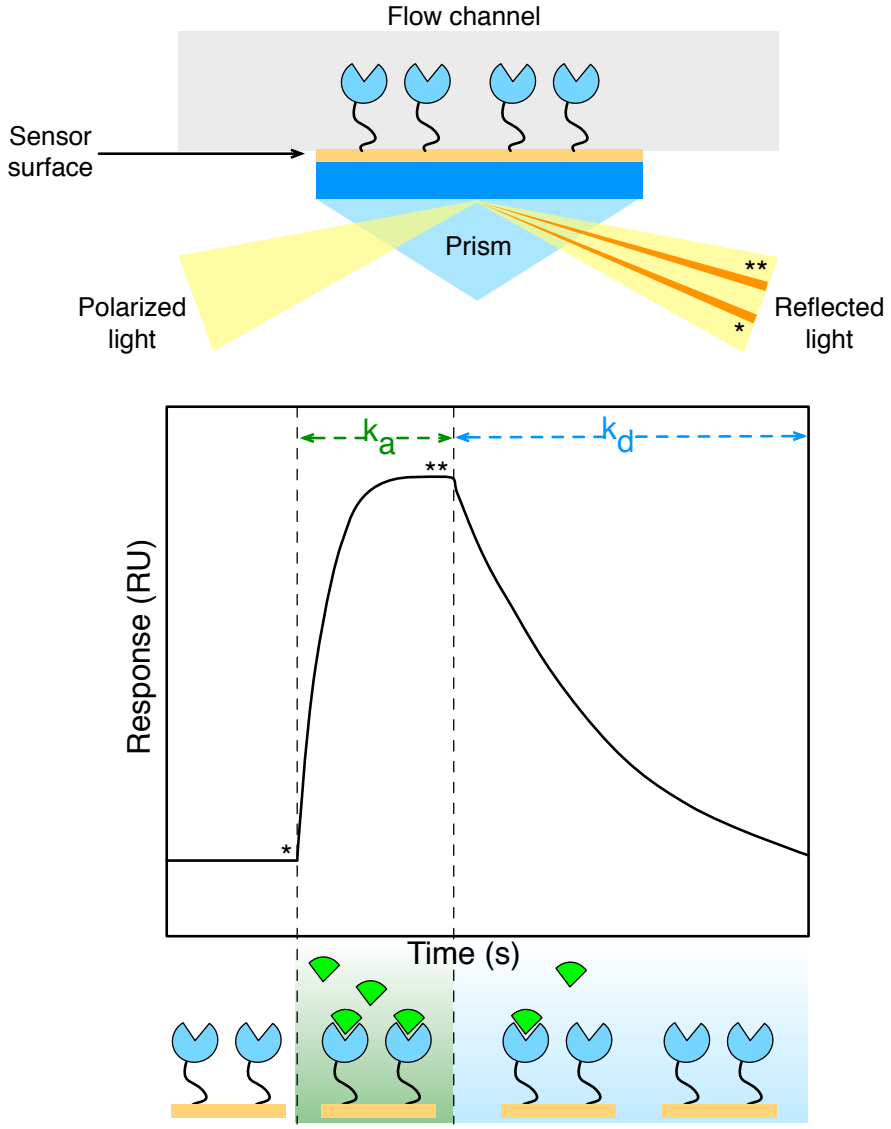



**Figure 3.2: Detection of binding events using SPR technology.**
Schematic of a surface plasmon resonance biosensor in the Kretschmann configuration (top). The p-polarized incident beam is directed at the sensor surface and the angle of incidence (SPR angle) is detected as a change in the detector position for the reflected intensity dip (*). Binding of an analyte to the immobilized ligand (1:1) at the sensor chip causes a corresponding increase in refractive index altering the SPR angle (**) and is shown as an increase in response in the sensorgram (bottom). Analyte dissociation results in a response drop as the SPR angle returns to the original position.

### 3.1.2.1 Sensor Chip

The Biacore CM sensor chip series is designed for efficient ligand immobilization to a surface suitable for SPR signal generation. It is composed of a glass surface coated with a thin gold layer (~ 50 nm thick) and covered with a flexible carboxymethylated (CM) dextran matrix. This matrix increases the immobilized ligand surface capacity, and is suitable for ligand attachment using various immobilization chemistries which exploit

ligand functional groups. Other derivitized matrices exist and the Biacore range also includes high affinity application sensor chips with streptavidin and nitrilotriacetic acid (NTA) for non-covalent capture of tagged ligands. However, covalent immobilization is often a preferred approach for ligand attachment as it avoids ligand removal during chip regeneration, provided suitable regeneration conditions can be found.

The CM matrix is suitable for a range of surface immobilization chemistries including thiol and aldehyde coupling. However, only amine coupling will be described here. Amine coupling involves activating the sensor chip surface with a mixture of 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide (EDC) and N-hydroxysuccinimide (NHS) prior to ligand attachment via primary amine groups. EDC activation of carboxylate molecules forms unstable *O*-acylisourea amine reactive intermediates, which are rapidly hydrolysed reforming the carboxyl group (Williams, Hill et al. 1981). Increasing the efficiency of EDC mediated amine coupling can be achieved with the addition of NHS, which converts the intermediate to a stable, amine-reactive succinimide ester (Gilles, Hudson et al. 1990) (Figure 3.3). Following the immobilization reaction, the sensor surface is treated with ethanolamine to deactivate excess reactive groups.

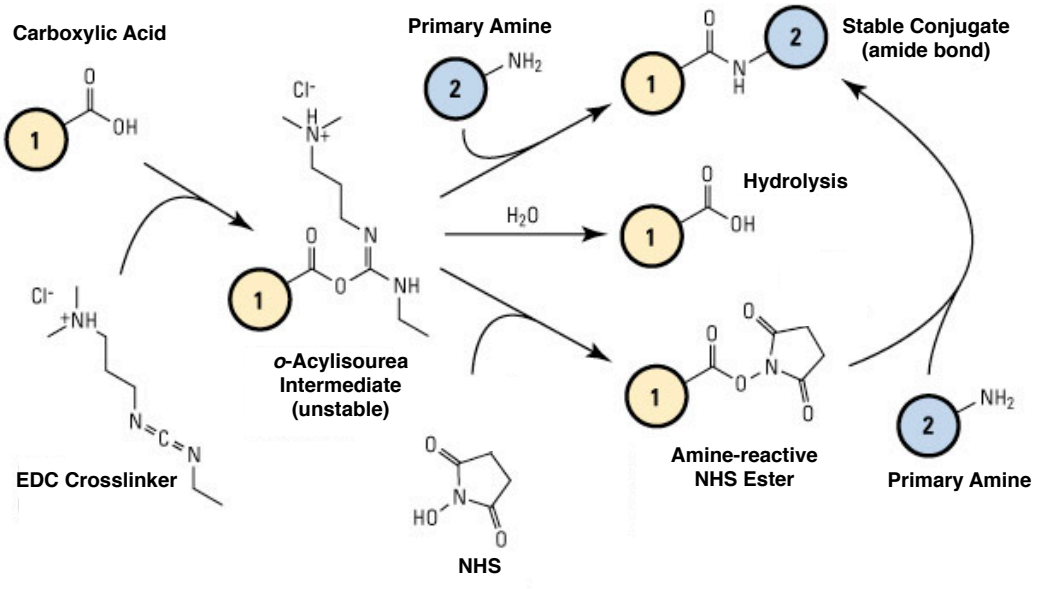



**Figure 3.3: NHS/EDC cross-linking reaction scheme.**
Carboxyl-to-amine cross-linking using the carbodiimide EDC and NHS. Molecules (1) and (2) can be peptides, proteins or any chemicals with carboxylate and primary amine groups. Addition of NHS to EDC reactions (*bottom pathway*) increases efficiency of EDC mediated amine coupling. (Image adapted from Carbodiimide Crosslinker Chemistry ThermoScientific 2009).

### 3.1.3 Calculation of binding affinity constants by SPR

Major applications of label-free real-time SPR technology include measurement of thermodynamic properties and binding rate constants of biospecific macromolecular interactions. Assuming a simple 1:1 interaction, the apparent rate and affinity constants are easy to calculate. The analyte (A) and immobilised ligand (B) associate at the sensor surface to form the complex (AB):

$$A+B \underset{k_d}{\overset{k_a}{\rightleftharpoons}} AB \tag{1}$$

and is defined by the association rate constant ($k_a$) and dissociation rate constant ($k_d$). The association phase is described by:

$$\frac{d[AB]}{dt} = k_a[A][B] - k_d[AB] \tag{2}$$

The amount of complex (AB) formed is proportional to the observed shift in SPR angle and is expressed in RU. At equilibrium,

$$\frac{d[AB]}{dt} = 0 \tag{3}$$

Therefore,

$$k_a[A][B] = k_d[AB] \tag{4}$$

Rearrangement of equation 4 gives the equilibrium constant ($K_d$), which describes the affinity of this interaction determined from the ratio of the kinetic rate constants, $k_d/k_a$, or from the level of binding at equilibrium as a function of sample concentration:

$$\frac{[A][B]}{[AB]} = \frac{k_d}{k_a} = K_d \tag{5}$$

Accurate interpretation of kinetic rate constants necessitates reduction of the mass transport effects, whereby the rate of ligand binding exceeds the rate of analyte transport to the sensor surface. This phenomenon results in an apparent association rate slower than that of the true association rate. Furthermore, rebinding artifacts during dissociation are exacerbated resulting in slower apparent dissociation rates. Low immobilisation levels (<100 RU) and high flow rates (>30 μL $min^{-1}$) are recommended to counteract mass transport effects (Myszka, Morton et al. 1997).

Preliminary experiments conducted by Dr Laurence Arnold indicated that the HPX-1:THP interaction was not amenable to analysis using kinetic rate constants due to rapid dissociation beyond the resolution limit of the instrumentation. Therefore, the SPR experimental data described in this chapter were analysed using equilibrium analysis:

$$R_{eq}=R_{max}\frac{[A]}{[A]+K_d} \quad (6)$$

where $R_{eq}$ is the response at equilibrium and $R_{max}$ is the maximum response.

### 3.1.4 The synthetic α1(I)772-787 Triple Helical Peptide

Collagen possesses unique physiochemical properties, which make it difficult to study. Conversely, small triple helical peptides (THPs) that mimic the structure of collagen are more amenable to study. Consequently, they have been used extensively in structural stability studies, protein-collagen binding interactions and as substrates and inhibitors for MMPs (Reviewed in Koide 2007, Fields 2010).

The 10.7 kDa homotrimeric THP used in this study is based on the α1(I) chain of type I collagen and has the sequence:

$$NH_2\text{-}C_6\text{-}(GPO)_4\,\underline{GPQG\sim IAGQRGVVGLPG}PO(GPO)_3\text{-}OH$$

in which 'O' denotes 4-hydroxyproline and the underlined 'guest' sequence is identical to residues Gly772-Gly787 of the collagen α1(I) chain. The THP, which was kindly donated by Prof. Gregg Fields (Torrey Pines Institute, Florida), has an N-terminal hexanoic group (designated $C_6$) followed by four GPO repeats, which together induce self-association and enhance thermal stability (Fields and Prockop 1996). The tilde (~)

marks the MMP-1 cleavage site within the THP 'guest' sequence, a position analogous to that in type I collagen (Lauer-Fields and Fields 2002, Chung, Yoshida et al. 2004).

## 3.2 Materials and Methods

The mutagenesis necessary to investigate the role of HPX-1 amino acid residues contributing to the putative collagen binding has been previously described in Chapter 2. Oligonucleotide design and preparation was undertaken as described in Section 2.3.4. Site-directed mutagenesis of target residues was completed as described in Section 2.3.5. Recombinant proteins were produced in accordance with protein expression and purification methods described in Table 2.5.

### 3.2.1 Preparation of the α1(I)772-787 Triple Helical Peptide

Prior to each experiment, the THP monomers were annealed into a triple helix at a high concentration (>1 mM) by heating for 15 min at 65 °C followed by slow cooling from 65 to 4 °C over 2 h. Because of their lack of aromatic residues, these α1(I)772–787 THP solutions were quantitated spectrophotometrically according to the formula: [THP] (mg/ml) = $((A_{215} - A_{225}) \times 0.144)/e$; where $e$ is the path length in cm (Sober 1968).

### 3.2.2 Surface Plasmon Resonance

All SPR experiments were performed using a Biacore T-100 instrument (GE Healthcare) at 25 °C.

#### 3.2.2.1 THP immobilisation to the CM4 sensorchip by amine-coupling

Ligand immobilisation was performed by Dr Andrew Pickford. In brief, the N-terminally aminated α1(I)772-787 THP was immobilized to the dextran surface of a CM4 sensorchip (GE Healthcare) flow cell using amine coupling in 10 mM HEPES pH 7.4. The more commonly used CM5 sensorchip (which has a higher degree of dextran carboxylation) was found to exhibit unacceptably high levels of non-specific electrostatic binding to WT HPX-1 (which has a pI of 9.45). The chip surface was activated prior to ligand immobilization using EDC and NHS solutions applied at the recommended flow rate and contact time specified in the GE Healthcare Biacore Sensor Surface Handbook (2008). Free succinimide esters were blocked with ethanolamine. All proprietary solutions were provided in the Amine-coupling Kit (GE Healthcare).

### 3.2.2.2 THP:HPX-1 Binding Assay

The running buffer for all SPR binding assays was HBSCP+ (10 mM HEPES, 150 mM NaCl, 10 mM $CaCl_2$, 0.05% (v/v) poly-sorbate-20), pH 7.4. Due to the extremely rapid dissociation rates of WT HPX-1 from immobilized α1(I)772-787 THP seen in preliminary experiments (which were beyond the T100 capability), the equilibrium dissociation constants ($K_d$s) for the wild-type and each mutant were determined from the SPR response at equilibrium. Thus, long-duration (10-min contact time) triplicate injections were performed with a suitable range of analyte concentrations (e.g. ~40 nM to ~40μM for WT HPX-1) at a flow rate of 1 μL/min. In each case, the equilibrium point was taken to be 5 s from the end of the analyte injection.

### 3.2.2.3 THP:HPX-1 Binding Data Analysis

Data analysis was performed using BIAevaluation™ software (GE Healthcare) by nonlinear least squares fitting using:

$$R_{eq}=R_{max}\frac{c}{c+K_d} \quad (7)$$

where $R_{eq}$ is the equilibrium response at a given concentration ($c$) of the WT or mutant protein, $R_{max}$ is the maximum response possible if the immobilized THP ligand were saturated with analyte, and $K_d$ is the dissociation constant.

### 3.2.3 Nuclear Magnetic Resonance of HPX-1 Mutants

1D $^{1}H$, and 2D $^{1}H$-$^{15}N$ heteronuclear single quantum coherence (HSQC) spectra of the purified isotopically $^{15}N$-labelled target proteins were acquired as described in Section 2.3.28 and 2.3.29.

## 3.3 Results

### 3.3.1 Production of HPX-1 mutants

Mutated pET-3a:HPX-1 constructs were successfully produced by Dr Laurence Arnold using site-directed mutagenesis and confirmed by DNA sequencing. These constructs were then transformed into *E. coli* expression strain BL21-CodonPlus (DE3)-RIPL and screened for IPTG-induced recombinant protein expression. A protein of ~24 kDa was expressed in all cases at levels equivalent to or even exceeding that of the WT HPX-1.

All target proteins were found to deposit into bacterial inclusion bodies and were therefore solubilised, refolded and purified in accordance with the standard procedures described in Chapter 2. However, during the refolding stage, three of the mutant HPX-1 domains (D299A, F316A and Q352A) precipitated leaving insufficient material for functional analysis. The remaining HPX-1 mutants that remained soluble during the refolding stage were found to elute from the cation exchange column at identical NaCl concentration to that of the WT HPX-1, suggesting that each protein possessed a native-like structure. In contrast, aggregation or an altered structural state unable to bind the structural $Ca^{2+}$ at the centre of the HPX β-propeller would change the charge properties of the recombinant proteins and therefore the subsequent elution profile.

Completion of the standard expression and purification procedure produced sufficient milligram quantities of target recombinant proteins sufficient for the assays described in this chapter.

### 3.3.2 Functional Analysis by SPR

Equilibrium analysis was used to characterise the binding affinity of WT HPX-1 and mutants thereof (R300A, F301A, S318A, V319A, P325A, R337A, D338A, Q354A, H358A, P361A) for immobilised α1(I)772-787 THP.

#### 3.3.2.1 Binding of WT HPX-1 to immobilised THP

Triplicate injections were performed using a suitable range of analyte concentrations (0.2 μM-45 μM) and affinities were determined by fitting to a 1:1 Langmuir isotherm binding model using the maximum response at equilibrium, taken to be 5 s from the end of the analyte injection. WT HPX-1 bound to α1(I)772-787 THP with an apparent equilibrium dissociation constant ($K_d$) value of 10.9 ± 0.9 μM (Figure 3.4). This finding opposes that of a previous study, which reported undetectable HPX-1 binding levels to collagenous heterotrimeric peptides (Ottl, Gabriel et al. 2000).

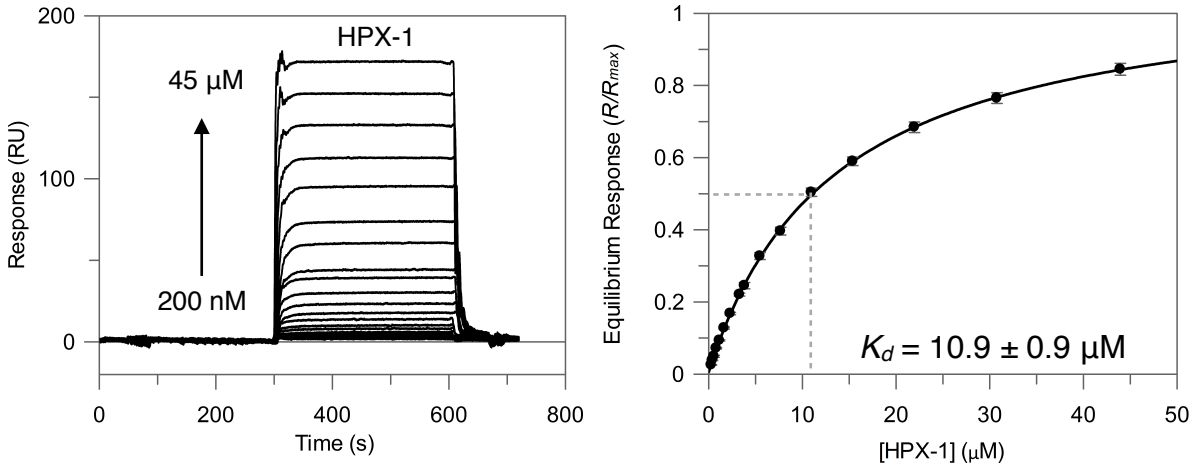
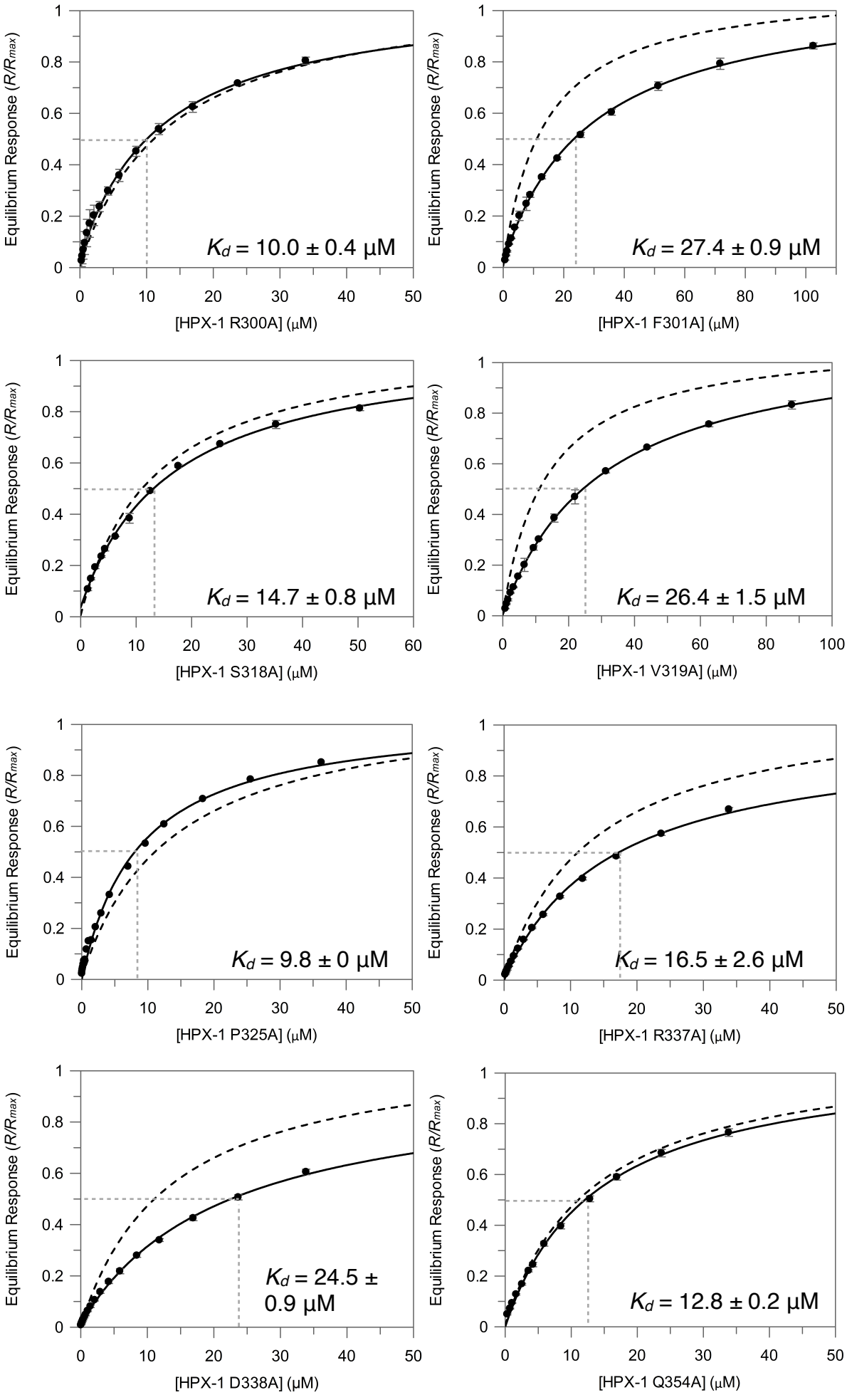



**Figure 3.4: Equilibrium binding of HPX-1 to α1(I)772-787 THP.** (**A**) Overlaid SPR sensorgrams showing the dose-dependent binding of the analyte, HPX-1, to an immobilised ligand, THP. (**B**) Hyperbolic binding curves for HPX-1 from the equilibrium SPR data. The fitted curve is of the form $R/R_{max} = (c/c + K_d)$, where $c$ is the analyte protein concentration. *Error bars* indicate the S.D. from triplicate measurements.

### 3.3.2.2 Screening mutant HPX-1 proteins

A minimum of duplicate injections were performed using a suitable range of analyte concentrations (~0.2 µM-~100 µM) and affinities were determined as described in Section 3.3.2.1.

HPX-1 mutants bound to α1(I)772-787 THP with apparent $K_d$ values between 9.8-27.4 µM (Figure 3.5). The binding activity of HPX-1 mutants R300A, S318A, P325A, Q354A and H358A were very similar to that of the WT HPX-1, suggesting that they play no direct role in contacting the collagen substrate. Conservative increases in $K_d$ were observed for mutants R337A and P361A. However, mutants F301A, V319A and D338A exhibited an approximate 3-fold reduction in binding relative to WT HPX-1 implying that these residues are important contributors to collagen recognition.


Equilibrium Response (R/Rmax)
[HPX-1 R300A] (μM)
Kd = 10.0 ± 0.4 μM
[HPX-1 F301A] (μM)
Kd = 27.4 ± 0.9 μM
[HPX-1 S318A] (μM)
Kd = 14.7 ± 0.8 μM
[HPX-1 V319A] (μM)
Kd = 26.4 ± 1.5 μM
[HPX-1 P325A] (μM)
Kd = 9.8 ± 0 μM
[HPX-1 R337A] (μM)
Kd = 16.5 ± 2.6 μM
[HPX-1 D338A] (μM)
Kd = 24.5 ± 0.9 μM
[HPX-1 Q354A] (μM)
Kd = 12.8 ± 0.2 μM

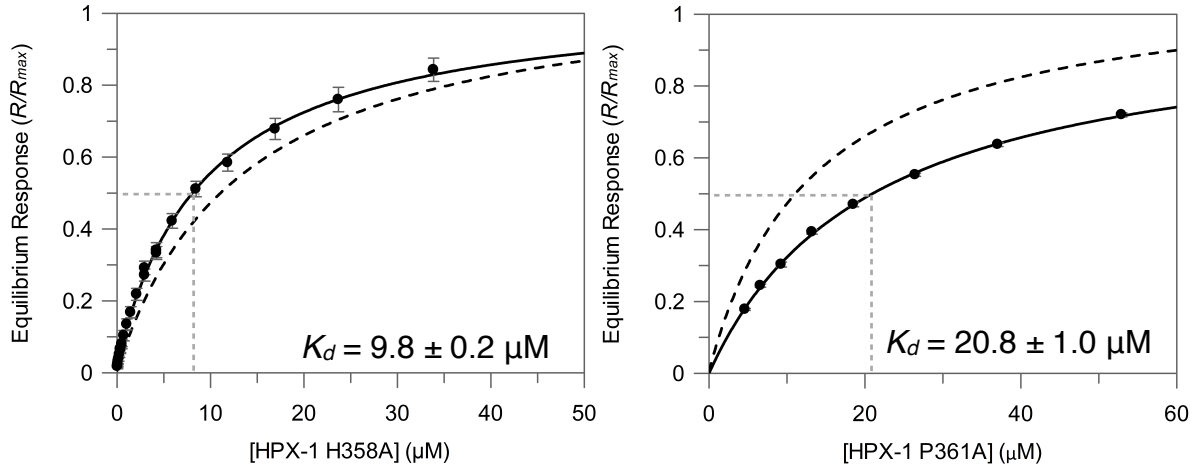



**Figure 3.5: Binding to the α1(I)772-787 THP by HPX-1 mutants.** Hyperbolic binding curves generated from the equilibrium SPR data. The fitted curve is of the form $R/R_{max} = (c/c + K_d)$, where $c$ is the analyte protein concentration. *Error bars* indicate the S.D. from replicate measurements. The *dashed line* shows concentration-dependent binding of wild-type HPX-1 for the THP substrate.

The binding affinities for all HPX-1 mutants tested are shown compared with WT HPX-1 in Figure 3.6.

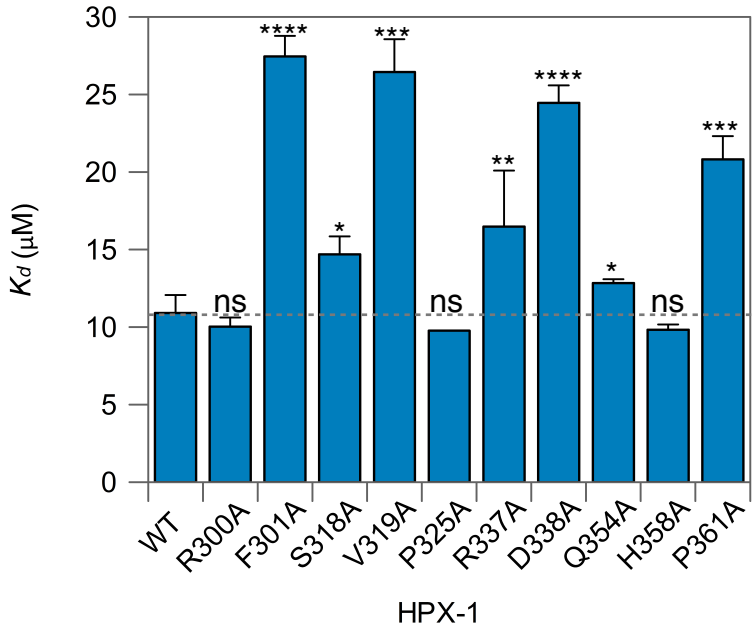



**Figure 3.6: Equilibrium dissociation constants ($K_d$) for the binding of WT HPX-1 and mutant proteins to immobilised THP.** *Error bars* indicate the S.D. from replicate measurements. The *horizontal dashed line* emphasises the $K_d$ of the WT HPX-1 protein for comparative purposes. *P* values were determined by comparing binding affinities of mutant proteins with WT HPX-1. Statistical analysis was performed using an unpaired two-tailed Student's *t*-test (df=4) and statistically significant *P* values ($p<0.05$) are shown. The asterisk indicates statistical significance with *P*-values of $p<0.05$ (*), $p<0.01$ (**), $p<0.001$ (***) and $p<0.0001$ (****). *P* values greater than or equal to 0.05 are not considered statistically significant (*ns*).

### 3.3.3 Structural integrity of HPX-1 F301A and V319A

Heteronuclear NMR spectroscopy was used to confirm that the reduced THP binding observed for HPX point mutants F301A and V319A was the result of selective removal of a critically important side-chain rather than perturbation of the native tertiary structure (Figure 3.7). In each case, the $^{1}H$, $^{15}N$-HSQC of the $^{15}N$-labeled HPX-1 mutant was very similar to that of the WT protein with excellent chemical shift dispersion in each dimension. In neither case were any extensive changes in $^{1}H$ or $^{15}N$ chemical shifts observed that would indicate a substantial change in the HPX-1 structure. Furthermore, no significant line broadening was observed in the HPX-1 mutants that could indicate a conformational exchange and thus destabilization of the domain.

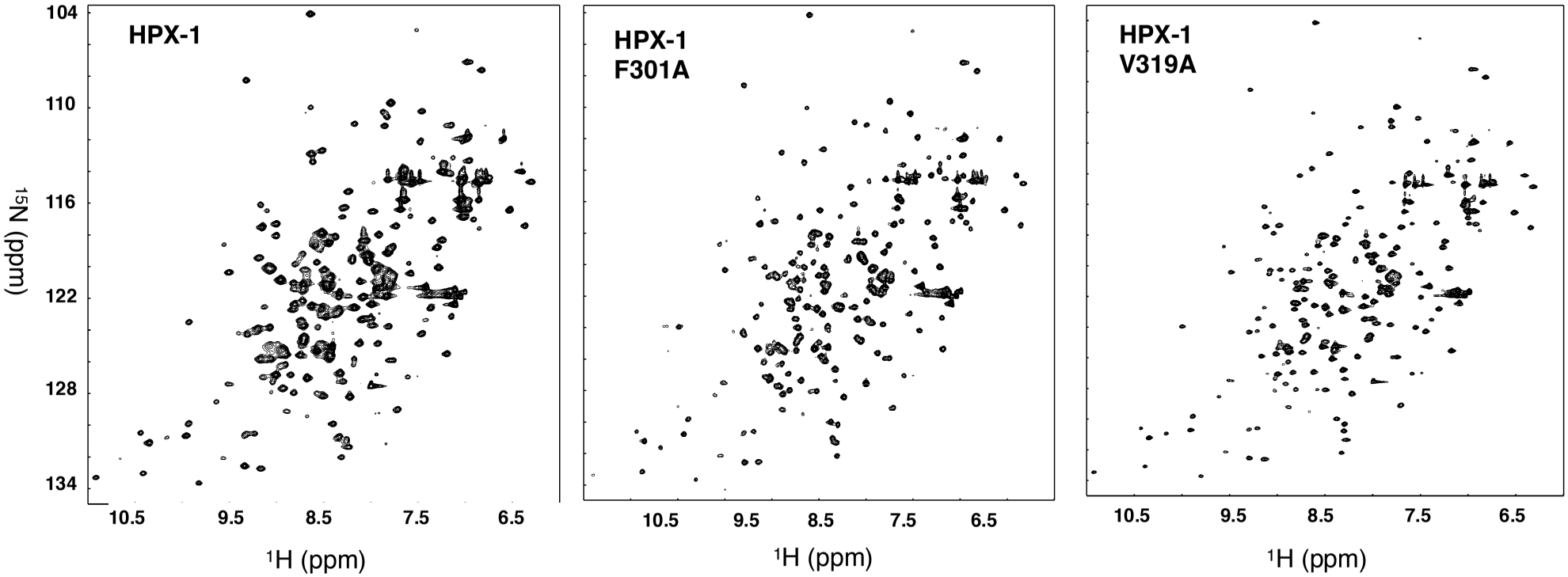



**Figure 3.7: $^{1}H^{15}N$- HSQC spectra of HPX-1 mutants F310A and V319A.** The spectra were recorded in 20 mM NaOAc pH 4.8, 5 mM $CaCl_2$, 50 µM $ZnCl_2$, 0.02% (w/v) $NaN_3$ and 10% $D_2O$ at 25°C on an Agilent Inova 600 MHz spectrophotometer. The $^{1}H^{15}N$- HSQC spectra of HPX-1 F301 A and V319A were recorded at concentrations of 170 µM and 173 µM, respectively. HPX-1 WT (283 µM) is shown for comparison.

## 3.4 Conclusions

Mutated pET-3a:HPX-1 constructs were successfully transformed into *E. coli* BL21-CodonPlus (DE3)-RIPL cells, prior to inducing protein expression. With the exception of D299A, F316A and Q352A HPX mutants, all proteins were successfully, refolded and purified at equivalent levels to WT HPX-1.

Despite a number of reports describing collagen binding by MMP HPX domains (Allan, Hembry et al. 1991, Murphy, Allan et al. 1992, Tam, Wu et al. 2002), a previous study observed no interaction between HPX-1 and a heterotrimeric THP using kinetic SPR analysis (Ottl, Gabriel et al. 2000). However, each chain of the $\alpha1(I)_2$ $\alpha2(I)$ THP used in the previous study lacked L785, part of the hydrophobic collagen region believed to contribute to MMP-1 recognition (Fields 1991, Leikina, Mertts et al. 2002). The synthetic THP utilised in this SPR-based assay incorporates both the MMP-1 cleavage site and L785, which may account for detectable HPX-1 binding thereby allowing elucidation of binding affinities ($K_d$ values).

All residues selected for testing were located within β-propellor blades 1 and 2 (Figure 3.8), implicated in collagen recognition during an NMR-monitored ligand titration. Of those residues tested using SPR, mutants that exhibited significantly decreased affinity (≥ 2 fold) for immobilised THP were F301A, V319A and D338A. Interestingly, D338 contributes to a salt-bridge with R291, a residue highlighted during analysis of MMP-1:THP interactions using HD/MS (Lauer-Fields, Chalmers et al. 2009), therefore alanine substitution at position 338 will disrupt the salt-bridge restricting ligand recognition. However, F301A and V319A HPX-1 mutations had the most pronounced effect on functional binding. $^{1}H^{15}N$- HSQC spectra of HPX-1 mutants F310A and V319A indicated that both proteins possessed folded tertiary structure, therefore the reduction in THP-binding for these mutants was taken as an indication of direct involvement in collagen recognition rather than deleterious structural effects from site-directed mutagenesis.

Intriguingly, F301 is concealed within the CAT-HPX domain interface as shown by examination of the mature MMP-1* crystal structure. Consequently, the role of this residue within the blade 1 exosite of the full-length enzyme remains to be elucidated. Further study is required and is the subject of Chapter 4.

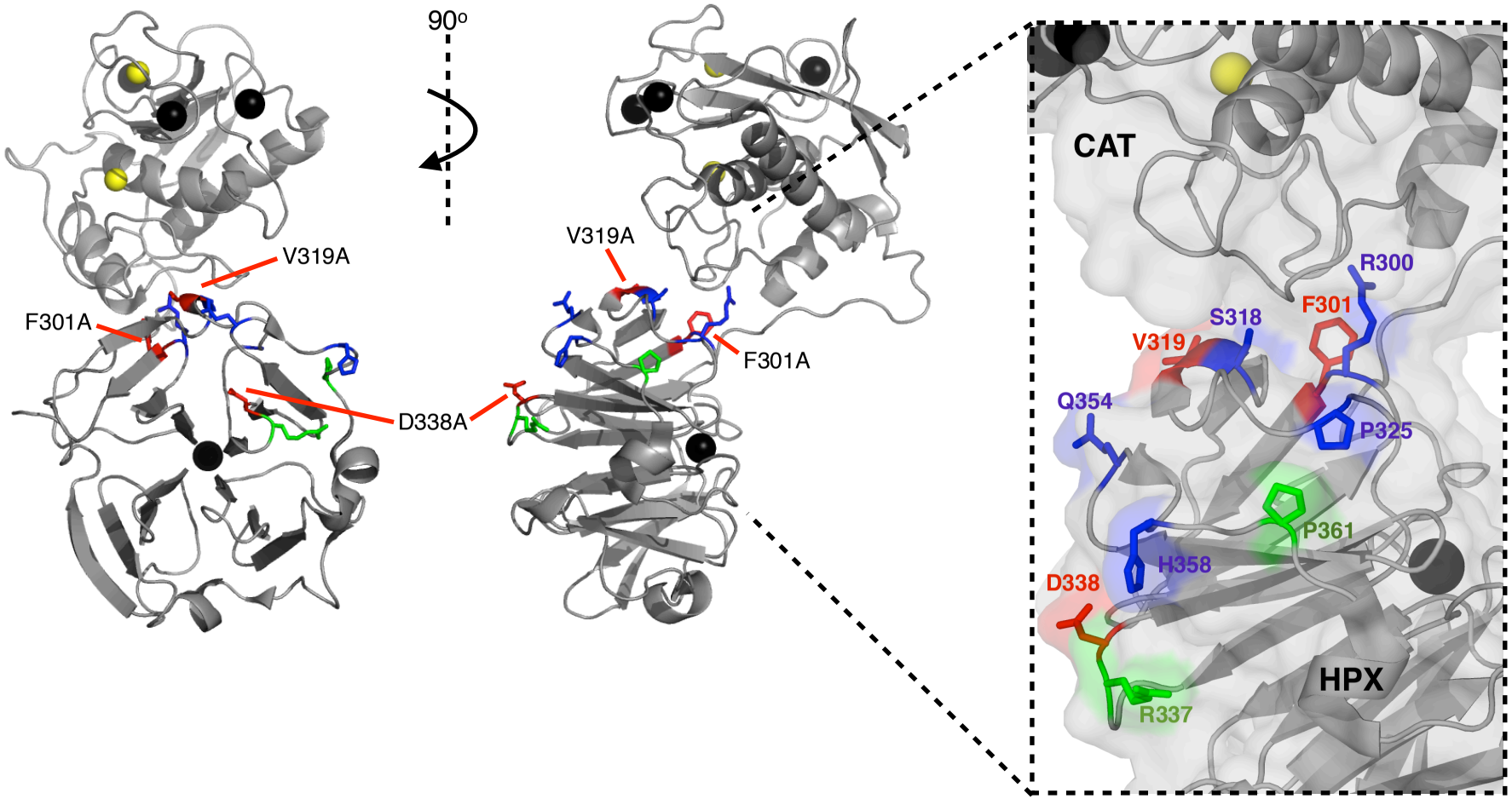



**Figure 3.8: Residues of HPX-1 implicated in collagen binding by mutagenesis and assay.** Orthogonal views of the crystal structure of MMP-1* are shown as a ribbon diagram. Mutated HPX-1 residues are shown as *sticks* and colored according to the collagen binding activity of the mutant, *i.e.* F301, V319, and D338 in *red* ($K_d \geq 24$ μM); R337 and P361 in *green* (16 μM < $K_d$ < 24 μM); and R300, S318, P325, Q354, and H358 in *blue* ($K_d \leq 16$ μM).

# CHAPTER 4: MMP-1 IS IN EQUILIBRIUM BETWEEN COMPACT AND DOMAIN-DISLOCATED STATES

## 4.1 Introduction

The lack of collagenolytic activity in HPX-deletion mutants of MMP-1, -8, -13 and -14 demonstrates the presence of collagenolytic determinants within the HPX domain (Hurst, Schwartz et al. 2004). These 'exosites' may, in part, explain how collagenase activity is achieved despite conflicting size dimensions of enzyme and substrate, which typically prevent peptide bond hydrolysis. It has been postulated that these exosites may impart collagenases with helicase activity, the ability to bind and partially unwind the triple helical collagen substrate (Tam, Moore et al. 2004), thereby allowing a single α-chain to enter the active site cleft of the CAT domain (Chung, Yoshida et al. 2004) and exposing the scissile bond necessary for sequential polypeptide cleavage. However, the molecular basis for this mechanism remains to be elucidated.

The NMR and SPR data obtained and presented in the previous chapter suggest that for MMP-1 an exosite exists in blade 1 of HPX-1. The highest impact mutation, F301A, reduced binding activity (2.5-fold) highlighting its role as an important contributor to collagen recognition. However, close inspection of the crystal structure reveals F301 lies at the end of β-strand D in blade 1 with its aromatic ring oriented away from the the core of HPX-1 and is buried at the CAT-HPX domain interface, a position likely to prohibit involvement in collagen binding in the full-length enzyme. Intriguingly, together with residues R300 and F316, residue F301 forms part of a ball and socket joint at the CAT-HPX domain interface (Figure 4.1). Rotation around this joint, may allow global structure reconfiguration of the compact arrangement, opening a groove between CAT-1 and HPX-1 that has been proposed to form the collagen binding site (Jozic, Bourenkov et al. 2005), a hypothesis that would explain the apparent lack of binding by proMMP-1 to collagen (Welgus, Jeffrey et al. 1981). Furthermore, differences in the reciprocal orientation of the CAT and HPX domains upon activation (discussed further in Chapter 6) and interdomain flexibility observed during solution studies of MMP-9 and -12 (Rosenblum, Cohen et al. 2007, Bertini, Calderone et al. 2008) support the notion that domain mobility may be a general feature of MMPs, and a significant property for the recognition and unwinding of triple helical collagen.

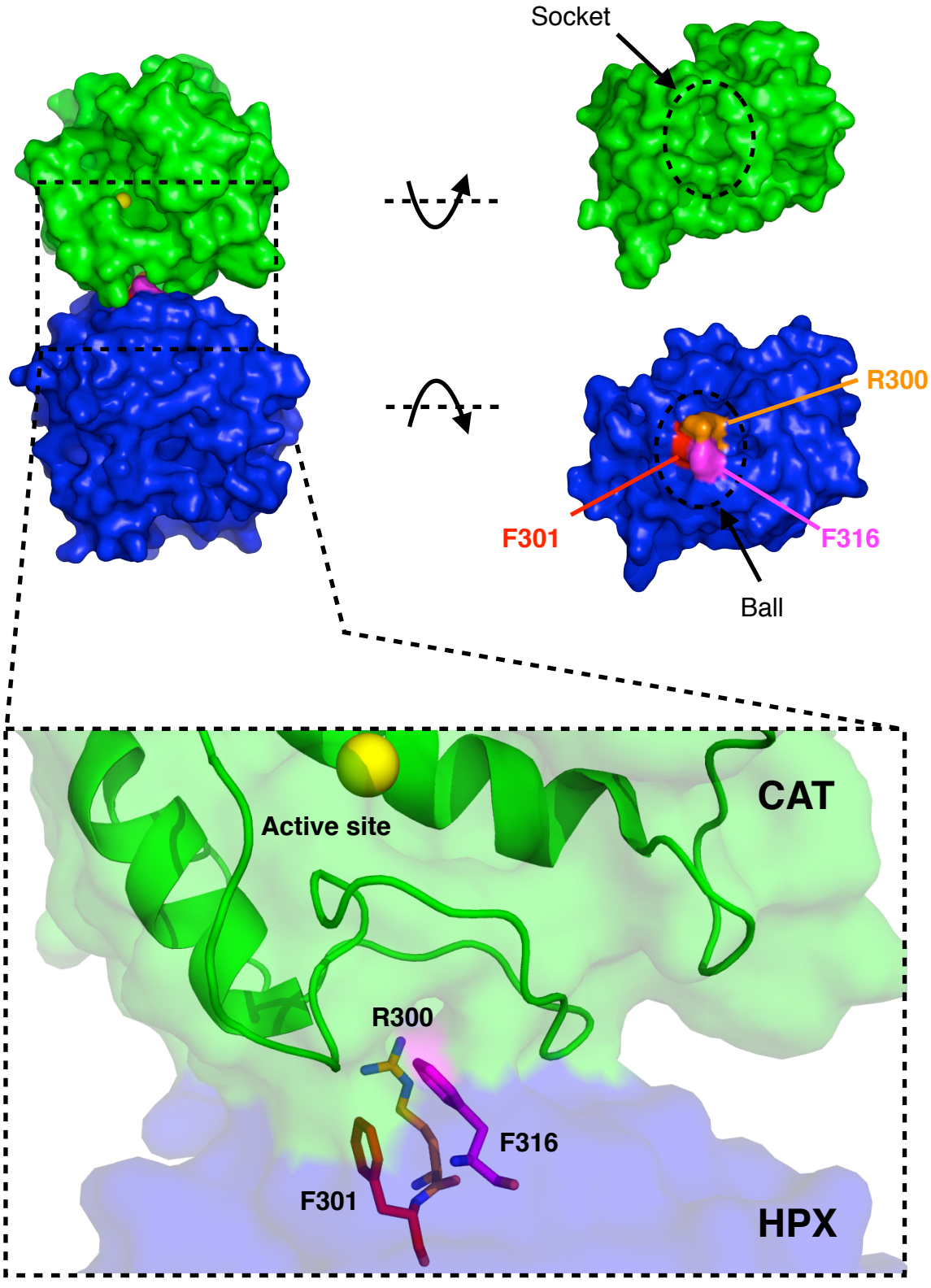



**Figure 4.1: The ball and socket joint in the CAT-HPX interface of MMP-1.** A space-filling model of the crystal structure of MMP-1 is shown looking into the active site cleft with the $Zn^{2+}$ ion colored *yellow*. Dislocation of the ball and socket joint and rotation by 90° of the CAT (*green*) and HPX (*blue*) domains reveals the interacting surfaces. The three residues comprising the ball are *highlighted*.

The study described in this chapter used alanine-scanning site-directed mutagenesis and assay to investigate the role of residues R300, F301 and F316. SPR was utilised to confirm the contribution of these residues towards binding a synthetic THP substrate, and a chromogenic assay was used to quantitate and characterise the catalytic activity of full-length mutant proteins. Finally, small angle x-ray scattering (SAXS) was used to investigate whether the compact arrangement of domains observed in the crystal structure is representative of the conformation in solution.

### 4.1.1 Small Angle X-Ray Scattering

Small angle x-ray scattering (SAXS) is a powerful technique that can be used for low-resolution structural characterization of biological macromolecules in solution and can provide information concerning the size, molecular weight, volume, shape and conformation of single molecules under a variety of conditions. The availability of high-brilliance synchrotron radiation sources, improving levels of automation during data acquisition and processing, and the ease of sample preparation have led to increasing popularity of this low-resolution structural technique (Reviewed in Svergun and Koch 2002, Lipfert and Doniach 2007). SAXS requires only micro-milligram quantities of the purified monodisperse target protein that can be monitored in varying conditions. In contrast, x-ray crystallography necessitates specialist crystallization conditions, while NMR demands isotopic-labelling of target proteins.

#### 4.1.1.1 SAXS Data Acquisition and Analysis

During a SAXS experiment monochromatic x-rays ($\lambda$ = 0.5-2Å) are used to irradiate a dilute solution of macromolecules (Figure 4.2). Attenuation of the incident radiation occurs as it interacts with the electron cloud of atoms within the macromolecule. At low angles of incidence, the beam ($k_i$) is elastically scattered, without energy transfer (Thomson scattering) and the synchronised plane waves produce interference patterns at the detector's position, the intensity of which depends on the orientation and distance of the scattering atoms in relation to the observation angle, $2\theta$.

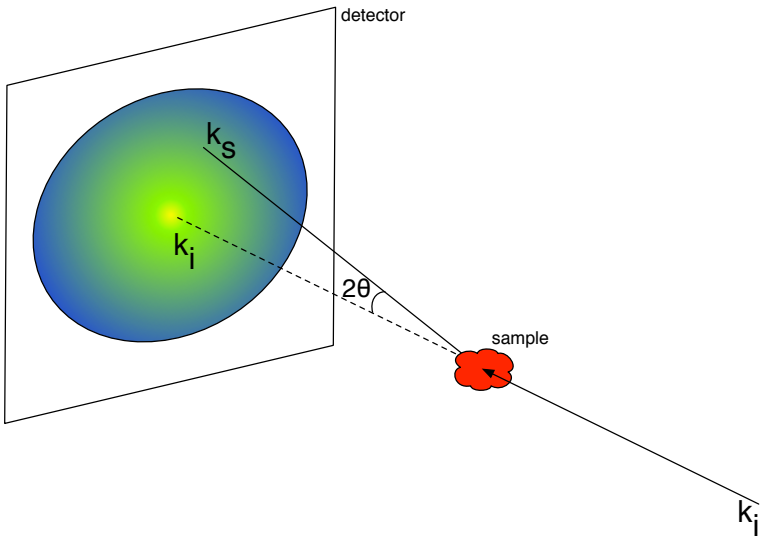



**Figure 4.2: Scattering geometry in small angle scattering.** Schematic showing the path of incidence ($k_i$) and scattered beams ($k_s$), respectively.

For elastic scattering both wave vectors have the magnitude $2\pi/\lambda$ and the scattered intensity (I), is measured as a function of the reciprocal space momentum transfer, $q$, which is defined as:

$$q = \frac{4\pi \sin\theta}{\lambda} \tag{1}$$

where $\lambda$ is the wavelength and 2θ is the scattering angle as shown in Figure 4.2. As such, the units of $q$ are inverse length units (Å$^{-1}$ or nm$^{-1}$). For randomly oriented scattering molecules, the two-dimensional intensity pattern is azimuthally averaged and resulting intensity $I(q)$ can be plotted as a function of the momentum transfer $q$ (Glatter and Kratky 1982, Feigin and Svergun 1987).

The amplitude of the scattered wave ($k_s$) depends on the electron density distribution $\rho(r)$ of the sample and therefore electron density contributions from both the scattering particle and the solvent must be considered (~0.33 e$^-$/Å$^3$ and ~0.44 e$^-$/Å$^3$, for pure water and protein respectively). The normalised scattered intensity after accurate solvent background subtraction is given by:

$$I(q) = NV^2 \Delta\rho^2 P(q) S(q) \tag{2}$$

where $N$ is the number density of scattering particles, $V$ is the average particle volume, $\Delta\rho$ is the scattering contrast between the particle of interest and the surrounding medium, $P(q)$ is the form factor describing the size and shape of the particles and $S(q)$ is

the structure factor describing particle interactions. For monodisperse dilute suspensions $S(q)\approx 1$, thus $I(q)$ is governed by the shape of the scattering molecules. In the case of non-interacting particles, $I(q)$ at small angles follows the Guinier approximation, which describes the moment of inertia for a particle using the electron density as a weighting factor:

$$I(q) = I(0)\exp\left(\frac{-R_g^{\,2}q^2}{3}\right) \tag{3}$$

where $I(0)$ is the intensity extrapolated to $q=0$ and $R_g$ is the radius of gyration. This is valid over a $q$ range limited by $qR_g<1.3$ for globular proteins (Putnam, Hammel et al. 2007). When (ln(I($q$)) versus $q^2$) is plotted the y-axis intercept provides ln($I(0)$) and the slope gives $-R_g^2/3$, from which the $R_g$ can be derived. Deviations from linearity in this region are indicative of sample aggregation, polydispersity or interparticle effects (Figure 4.3).

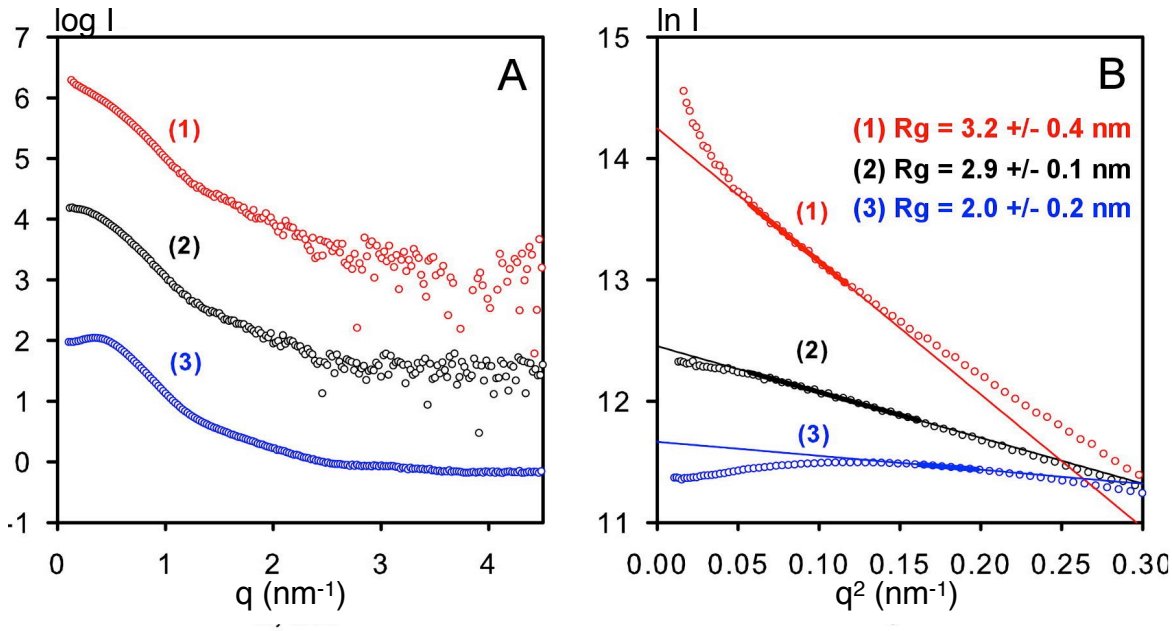



**Figure 4.3: Guinier Analysis.** Standard SAXS curves (**A**) and Guinier plots (**B**) for BSA samples measured at X33 (DORIS, Hamburg) in different buffers showing (1) aggregation, (2) good data and (3) inter-particle repulsion. The Guinier fits for estimation of $R_g$ and $I(0)$ are displayed, with the linear regions used for parameter estimation indicated by the thick lines (Adapted from Mertens and Svergun 2010).

Inversion of reciprocal space using a Fourier transform of the scattered intensity, I($q$) is known as the pair distance distribution function, *P(r)*:

$$I(q) = 4\pi \int_0^\infty P(r) \frac{sin(qr)}{q(r)} dr \tag{4}$$

which describes the paired intramolecular atomic distances of a particle and is useful for detecting overall shape, conformational changes (Figure 4.4), real space approximation of $R_g$ and maximal dimensions ($D_{max}$), which have the smallest frequency and can therefore be determined by the x-axis intercept (Feigin and Svergun 1987). The function requires measurement over the entire $q$ range, which is prevented as scattering rapidly decreases at high $q$ angles. Typically, the *P(r)* function is therefore calculated by indirect Fourier transformation to avoid the inherent problems of sampling over a finite $q$ range (Glatter 1977).

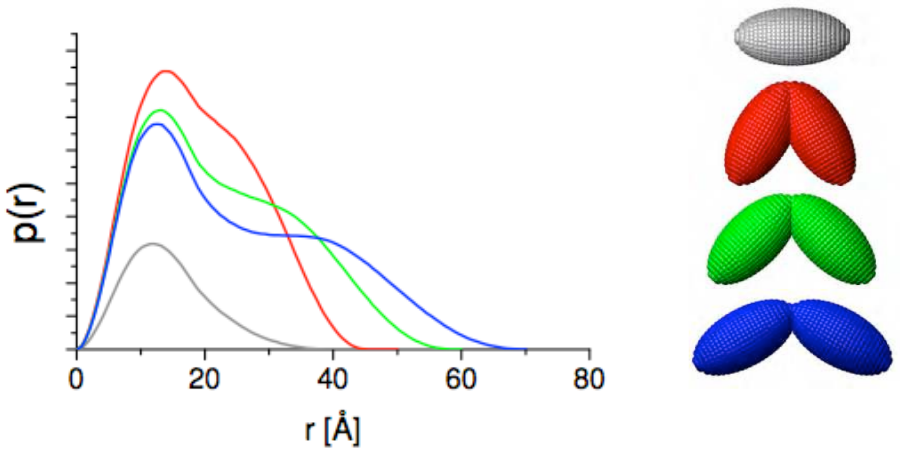



**Figure 4.4: Theoretical *P(r)* functions for monomeric and dimeric bodies with simple shapes.** The bead models used for the calculation are shown. Conformational variation is easily detectable by observing changes in peak height, shape and maximal dimension (Reproduced from Chandrasekaran 2011).

### 4.1.1.2 SAXS Data Processing and Analysis Software

The quantity of data generated by small angle scattering experiments continues to increase as a result of hardware developments. Large area hybrid pixel detectors developed for high-energy experiments feature high dynamic ranges, millisecond read-out times and single pixel photon counting (Broennimann, Eikenberry et al. 2006) with

count rates of $3{\cdot}10^8$ photons per $mm^2$ per sec recorded (Dectris® 2012). Therefore, users can easily expect to generate several gigabytes of data during a single SAXS visit.

To manage SAXS experimental data evaluation several program packages are freely available. Developed at EMBL, Hamburg, ATSAS (Konarev, Petoukhov et al. 2006) is one popular program suite dedicated to biological macromolecule analysis and allows users to accomplish most data analysis tasks. The workflow begins with raw data processing and finishes with a low-resolution three-dimensional (3D) structural model. PRIMUS (Konarev, Volkov et al. 2003) is a graphical package for preliminary analysis of 1D data. Using this programme, it is possible to complete a variety of data manipulations including buffer subtraction, merging data from different angular ranges, extrapolation to infinite dilution, and evaluation of key SAXS parameters including the $R_g$ and $I(0)$ . Good quality data can be further analysed using GNOM (Svergun 1992), an indirect Fourier Transform programme, that uses the 1D scattering data to calculate the distance distribution function for the target macromolecule. Successive iterations are required to specify an appropriate distribution range and optimum $D_{max}$ value; this parameter is user-defined and so care must be taken to ensure a good solution is found.

Three dimensional *ab initio* reconstructions can be derived from 1D scattering intensity data and several shape restoration programmes are available for determining macromolecular shapes. DAMMIN and DAMMIF use the GNOM output file to produce low-resolution structures. (Volkov and Svergun 2003, Franke and Svergun 2009) Using simulated annealing of dummy atoms in a spherical search volume, the programs are able to calculate optimal protein configurations that fit the experimental SAXS data.

Known high-resolution atomic structures can be compared to experimental intensity data using CRYSOL (Svergun, Barberato et al. 1995), which generates a theoretical SAXS scattering profile for the known structure and uses the $chi^2$ agreement to evaluate the fit. The *ab initio* envelopes can be superimposed using SUPCOMB (Kozin and Svergun 2001), a program that minimises the normalised spatial discrepancy (NSD) to find the best alignment between models. Model alignment discrepancy may be apparent for proteins with disordered regions or proteins experiencing conformational freedom.

## 4.2. Materials and Method

The mutagenesis and SPR methods necessary to investigate the role of residues Arg300, Phe301 and Phe316 have previously been described in Chapters 2-3. Oligonucleotide design and preparation was undertaken as described in Section 2.3.4. Site-directed mutagenesis of target residues was completed as described in Section 2.3.5. Recombinant proteins were produced in accordance with protein expression and purification methods described in Section 2.3. Substrate-binding by proMMP-1* and MMP-1* recombinant proteins was assayed using SPR as described in Section 3.2.2. Substrate cleavage by active MMP-1 recombinant proteins was conducted as described in Section 2.3.23.

### 4.2.1 Analytical Gel Filtration

Purified samples were concentrated, using centrifugal filter units with 10 kDa MWCO (Millipore), to a volume of 1 mL using a Beckman Coulter Allegra 25R refrigerated centrifuge (rotor AT-14-10) pre-chilled to 4 $^{o}$C. Samples were resuspended after each spin (3000 rpm for 1-10 minutes). Following concentration, the samples were applied at 1 mL/min to a HiLoad Superdex 75 16/60 prep grade column pre-equilibrated with TNC buffer. 1 mL fractions were collected and were analysed by SDS-PAGE (Section 2.3.19.

### 4.2.2 Small Angle X-Ray Scattering Data Acquisition

Recombinant proMMP-1* and MMP-1* proteins purified by gel filtration (Section 2.3.22) were dialysed overnight (Spectrum Standard Regenerated Cellulose Spectra/Por dialysis membrane MWCO 3.5 kDa) against 1000 volumes of TNC buffer at 4 $^{o}$C prior to transportation on ice to the Diamond Light Source (DLS), Harwell Science and Innovation Campus, Didcot. Protein solutions were concentrated as described in Section 2.3.22 and the protein concentration was assessed. SAXS data was collected on the high-brilliance I22 beamline using a wavelength ($\lambda$) of 1.0 Å and an energy of 12.4 keV, and equipped with a photon counting detector at a distance of 3.25 m from the sample cell. The scattering patterns were measured for several protein concentrations in the range from 0.84 to 8.25 mg/ml, each with multiple successive 1s exposures over a 60s period (to check for radiation-induced damage and aggregation). Bovine serum albumin (BSA) was used for calibration purposes. Measurements of the dialysis buffer

at equilibrium with the sample were used for buffer subtraction and to assess beam intensity fluctuations.

### 4.2.3 Small Angle X-Ray Scattering Data Processing and Analysis

The data were normalised to the intensity of the beam, radially averaged and buffer scattering was subtracted using the Non-crystalline Diffraction data reduction programme provided by DLS. Successive frames displaying changes in scattering intensity, indicative of radiation damage, were removed before further processing. A $q$ range of between 0.016 and 0.609 $Å^{-1}$ was used in the analysis to exclude the parasitic scattering near the beam stop. Data were analysed and processed using the ATSAS software package (Konarev, Petoukhov et al. 2006). The data were extrapolated to infinite dilution using the program PRIMUS (Konarev, Volkov et al. 2003) and the forward scattering, $I(0)$, and the radius of gyration, $R_g$, were evaluated using the Guinier approximation, assuming that at very small angles ($s < 1.3/R_g$).

The molecular mass of proMMP-1* was estimated using the following equation:

$$MM_p = I(0)_p / c_p \frac{MM_{st}}{I(0)_{st} / c_{st}} \quad (5)$$

where $I(0)_p$ and $I(0)_{st}$ are the scattering intensities at zero angle of the target protein and the standard protein, respectively. $MM_p$ and $MM_{st}$ are the corresponding molecular masses and $c_p$ and $c_{st}$ are the relevant concentrations (Mylonas and Svergun 2007). The $P(r)$ function was used to ensure a suitable number of data points was included for extrapolation of accurate I(0) values for both the target protein and the BSA standard (Putnam, Hammel et al. 2007).

The values of $I(0)$, the real space $R_g$, the maximum dimension ($D_{max}$), and the interatomic distance distribution function, $P(r)$, were calculated using the program GNOM (Svergun 1992). Simulated SAXS curves were generated from the atomic resolution crystal structures of proMMP-1* and MMP-1* (PDB accession codes 1SU3 and 2CLT, respectively) using the program CRYSOL (Svergun, Barberato et al. 1995). Where appropriate, ten independent *ab initio* low resolution atomic reconstructions were generated from the $P(r)$ data by simulated annealing using the DAMMIF program (Franke and Svergun 2009), which was run using default parameters and no enforced

particle symmetry. DAMMIF *ab initio* models were individually checked for outliers and then aligned, averaged, and filtered using DAMAVER (Volkov and Svergun 2003). Overlay of the SAXS models with those from x-ray crystallography was performed with the program SUPCOMB (Kozin and Svergun 2001), in which a low NSD value indicated an acceptable fit.

## 4.3. Results

### 4.3.1 Mutagenesis

Site-directed mutagenesis of proMMP-1* and proMMP-1 plasmids was completed using the Quikchange® II XL Site-Directed Mutagenesis Kit (Agilent Technologies). Mutagenic efficiency was monitored by blue/white colour selection of transformed pWhitescript control mutagenesis reactions. Mutagenic efficiency was calculated as the percentage of colonies displaying a blue phenotype indicative of incorporation of the functional (*lac*$Z^+$) mutation. Typical mutagenic efficiency levels were in excess of 94%. Transformations of mutated proMMP-1* and proMMP-1 plasmids yielded 10-150 cfu/ 250 µL of transformation reaction. Clones potentially harbouring mutants plasmids were randomly selected and harvested using a QIAprep Spin Miniprep Kit prior to dilution to 100 ng/µL in preparation for sequencing. DNA sequencing was undertaken by the QuickLane Express Sequencing service provided by Beckman Coulter Genomics. DNA sequence analysis confirmed successful incorporation of mutagenic bases and hence single alanine amino acid substitutions in either proMMP-1* or proMMP-1 (See Appendix).

### 4.3.2 Protein Production of Zymogens

Successfully mutated proMMP-1* and proMMP-1 plasmids were transformed into *E. coli* expression strain BL21-CodonPlus® (DE3)-RIPL and screened for IPTG-induced recombinant protein expression. A protein of approximately 52 kDa was expressed in all cases, this is consistent with the expected mass predicted for proMMP-1 using the ExPASy ProtParam tool (Gasteiger, Gattiker et al. 2003) (See Appendix).

All target proteins were found to deposit into bacterial inclusion bodies and were therefore solubilised, refolded and purified in accordance with standard procedure (See Chapter 2). Soluble target proteins were found to elute from the cation exchange column at an identical NaCl concentration to that of proMMP-1*, suggesting that each protein possessed a native-like structure. In contrast, aggregation or an altered structural state unable to bind the structural $Ca^{2+}$ at the centre of the HPX β-propeller would change the charge properties of the recombinant proteins and therefore the subsequent elution profile.

Completion of the standard expression and purification procedure produced sufficient milligram quantities of monodisperse target recombinant proteins for the assays described in this chapter (Table 4.1).

Table 4.1: Purification of proMMP-1* proteins[ab]

| Protein | Fraction | Total Volume (mL) | Protein Concentration (mg/mL) | Total Protein (mg) | Recovered Protein (%) |
|---|---|---|---|---|---|
| proMMP-1* | Solubilised Inclusion Bodies | 20 | 4.46 | 89.2 | 100 |
| | Post-CEX | 30 | 0.97 | 29.1 | 32.62 |
| | Post-GF | 25 | 0.60 | 15.0 | 16.82 |
| proMMP-1* R300A | Solubilised Inclusion Bodies | 20 | 7.87 | 157.4 | 100 |
| | Post-CEX | 30 | 0.44 | 13.2 | 8.38 |
| | Post-GF | 25 | 0.24 | 5.99 | 3.81 |
| proMMP-1* F301A | Solubilised Inclusion Bodies | 20 | 5.88 | 117.6 | 100 |
| | Post-CEX | 30 | 0.48 | 14.4 | 12.24 |
| | Post-GF | 25 | 0.23 | 5.75 | 4.89 |
| proMMP-1* F316A | Solubilised Inclusion Bodies | 20 | 5.45 | 109.0 | 100 |
| | Post-CEX | 30 | 0.41 | 12.3 | 11.28 |
| | Post-GF | 25 | 0.23 | 5.63 | 5.16 |

[a] from 500 mL of culture
b protein concentration was determined spectrophotometrically ($A_{280}$) using the molar extinction coefficient values obtained from ProtParam analysis (Gasteiger, Gattiker et al. 2003).

#### 4.3.2.1 Protein Production of Mature Enzymes

Recombinant proMMP-1* and proMMP-1 proteins were treated with APMA and CAT-3, and all showed complete enzyme maturation within a four hour incubation period concomitant with an appropriate reduction in molecular weight of ~10 kDa as shown by SDS-PAGE analysis (data not shown). During the final purification by gel filtration chromatography, it was noted that 'ball' mutant proteins appeared to elute earlier than expected suggestive of a less compact structure in solution.

### 4.3.3 Analytical Gel Filtration

Analytical gel filtration chromatography was used to maximise the resolution of the elution profile of activated proteins. All three mutants exhibited a decrease in elution

volume, which correlated with an increase in hydrodynamic radius relative to MMP-1* (Figure 4.5) and is suggestive of a less compact structure.

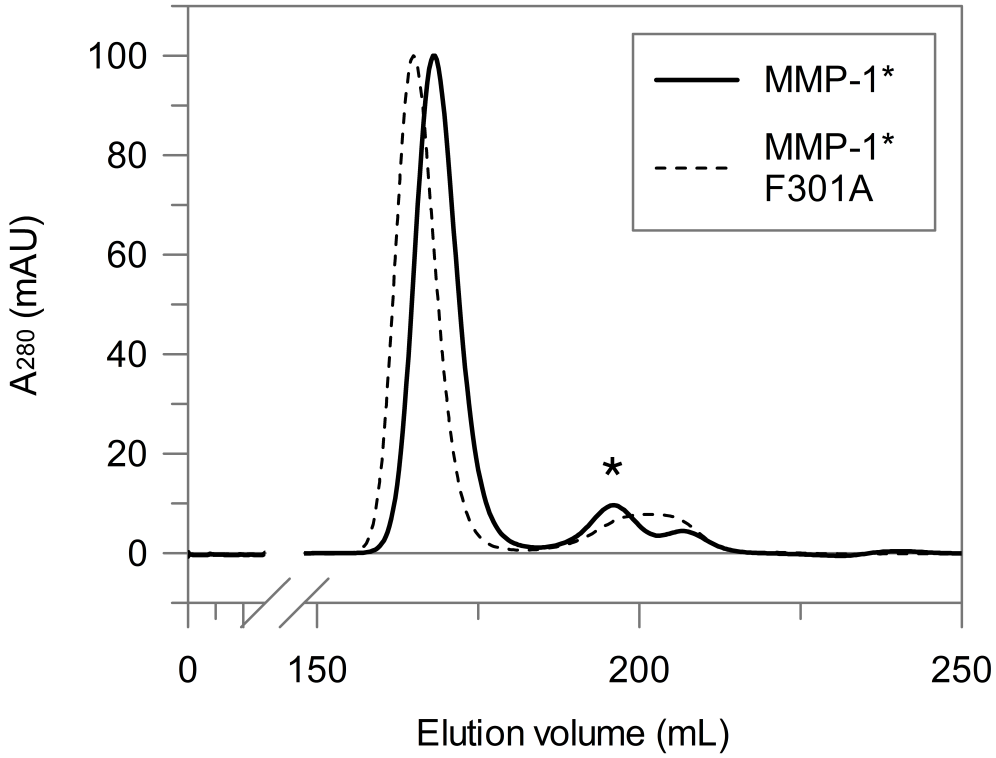



**Figure 4.5: Gel filtration of MMP-1* and MMP-1* F301A.** Mutation of 'ball' residues induces conversion to a less compact form as shown by the reduced elution volume. Elution of CAT-3 is shown by an *asterisk* (*). Elution profiles for R300A and F316A were superimposable upon that for F301A but have been omitted for clarity.

### 4.3.4 Functional Analysis by SPR

Equilibrium SPR analysis was used to characterise the binding affinity of the proMMP-1* zymogen, mature MMP-1* and the MMP-1* 'ball and socket' mutants (R300A, F301A and F316A) for immobilised α1(I)772-787 THP.

Triplicate injections were performed using a suitable range of analyte concentrations (~0.1 nM-40 μM) and affinities were determined by fitting to a 1:1 Langmuir isotherm binding model using the maximum response at equilibrium, taken to be 5 s from the end of the analyte injection.

### 4.3.4.1 Binding of proMMP-1* to Immobilised THP

The zymogen, proMMP-1* bound to α1(I)772-787 THP with an apparent equilibrium dissociation constant ($K_d$) value of 7.37 ± 1.34 μM (Figure 4.6). Interestingly, this is in contrast to previous reports, which claim that proMMP-1 lacks the capacity to bind collagen (Welgus, Jeffrey et al. 1985, Murphy, Allan et al. 1992).

The experimentally-derived $K_d$ value is similar to that observed for HPX-1 alone ($K_d$ = 10.9 ± 1.0 μM) suggesting that binding is most likely solely mediated by contributions from the HPX domain. Occlusion of the CAT domain active site by the inhibitory propeptide would support this observation.

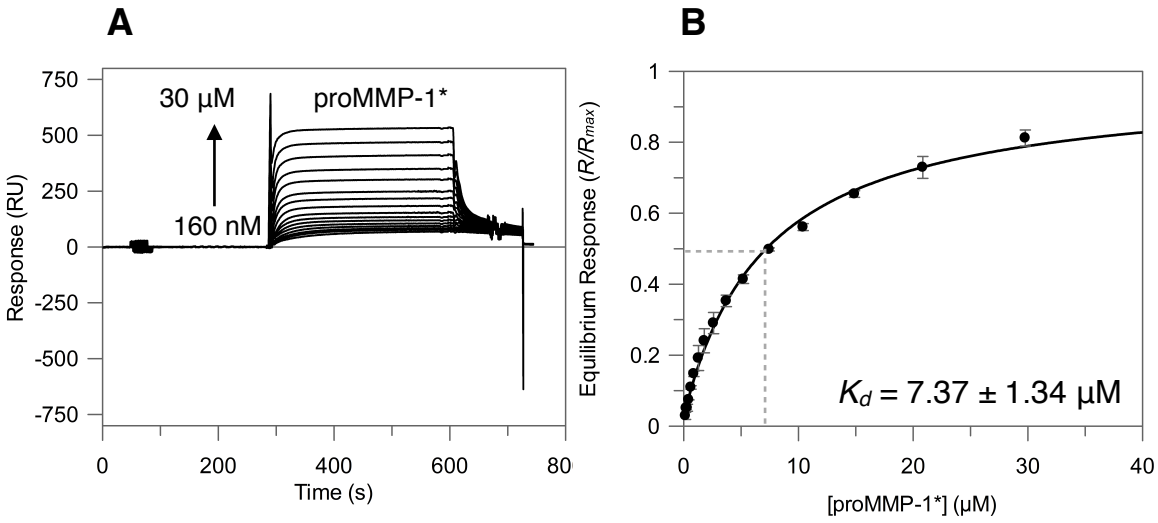



**Figure 4.6: Equilibrium binding of proMMP-1* to α1(I)772-787 THP.** (**A**) Overlaid SPR sensorgrams showing the dose-dependent binding of the analyte, proMMP-1*, to an immobilised ligand, THP. (**B**) Hyperbolic binding curve for proMMP-1* from the equilibrium SPR data. The fitted curve is of the form $R/R_{max} = c/(c + K_d)$, where $c$ is the analyte protein concentration. Error bars indicate the S.D. from triplicate measurements.

### 4.3.4.2 Binding of MMP-1* to Immobilised THP

MMP-1* bound to α1(I)772-787 THP with an apparent $K_d$ value of 0.92 ± 0.09 μM (Figure 4.7), an order of magnitude lower than HPX-1 alone ($K_d$ = 10.9 ± 1.0 μM). Therefore, it is clear that the observed increase in affinity by the full-length enzyme is a consequence of contributions made by the CAT domain and/or the CAT-HPX linker.

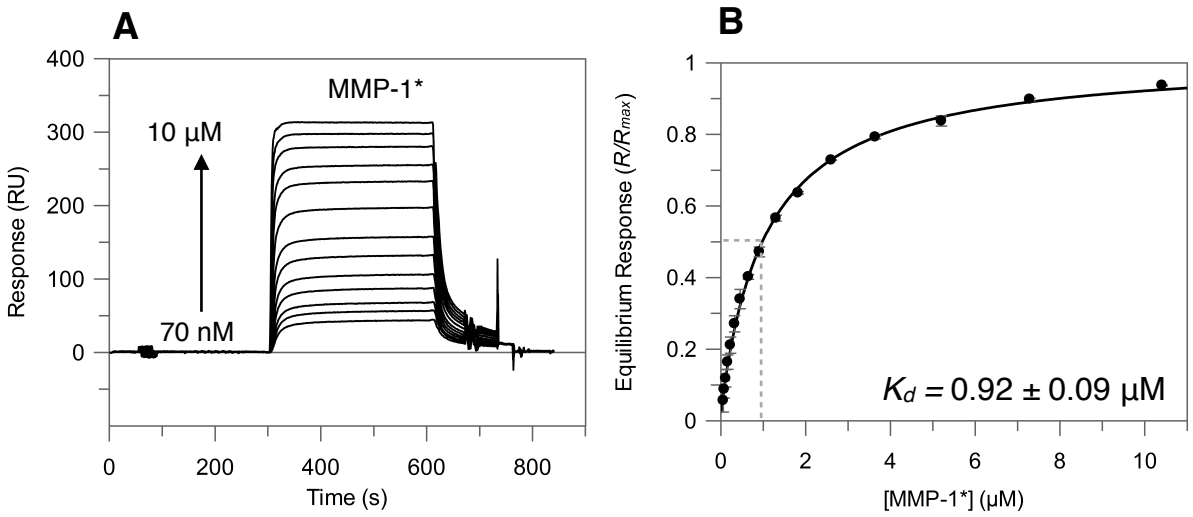



**Figure 4.7: Equilibrium binding of MMP-1* to α1(I)772-787 THP.** (**A**) Overlaid SPR sensorgrams showing the dose-dependent binding of the analyte, MMP-1*, to an immobilised ligand, THP. (**B**) Hyperbolic binding curve for MMP-1* from the equilibrium SPR data. The fitted curve is of the form $R/R_{max} = c/(c + K_d)$, where $c$ is the analyte protein concentration. Error bars indicate the S.D. from triplicate measurements.

### 4.3.4.3 Binding of Mutant proMMP-1* Proteins (R300A and F316A) to Immobilised THP

Zymogen proteins, proMMP-1* R300A and F316A, bound to α1(I)772-787 THP with apparent $K_d$ values of 10.68 ± 0.40 µM and 7.29 ± 0.11, respectively (Figure 4.8). Interestingly, the F316A mutation appeared not to affect THP binding; in contrast the R300A mutation had an adverse effect, reducing binding by 45%. However, statistical analysis revealed no significant difference between the mutants and proMMP-1*.

It is anticipated that the impact of the F301A mutation in the weak-binding zymogen form would require high protein concentrations likely to impede elucidation of $K_d$ values, therefore this particular mutation remains untested.

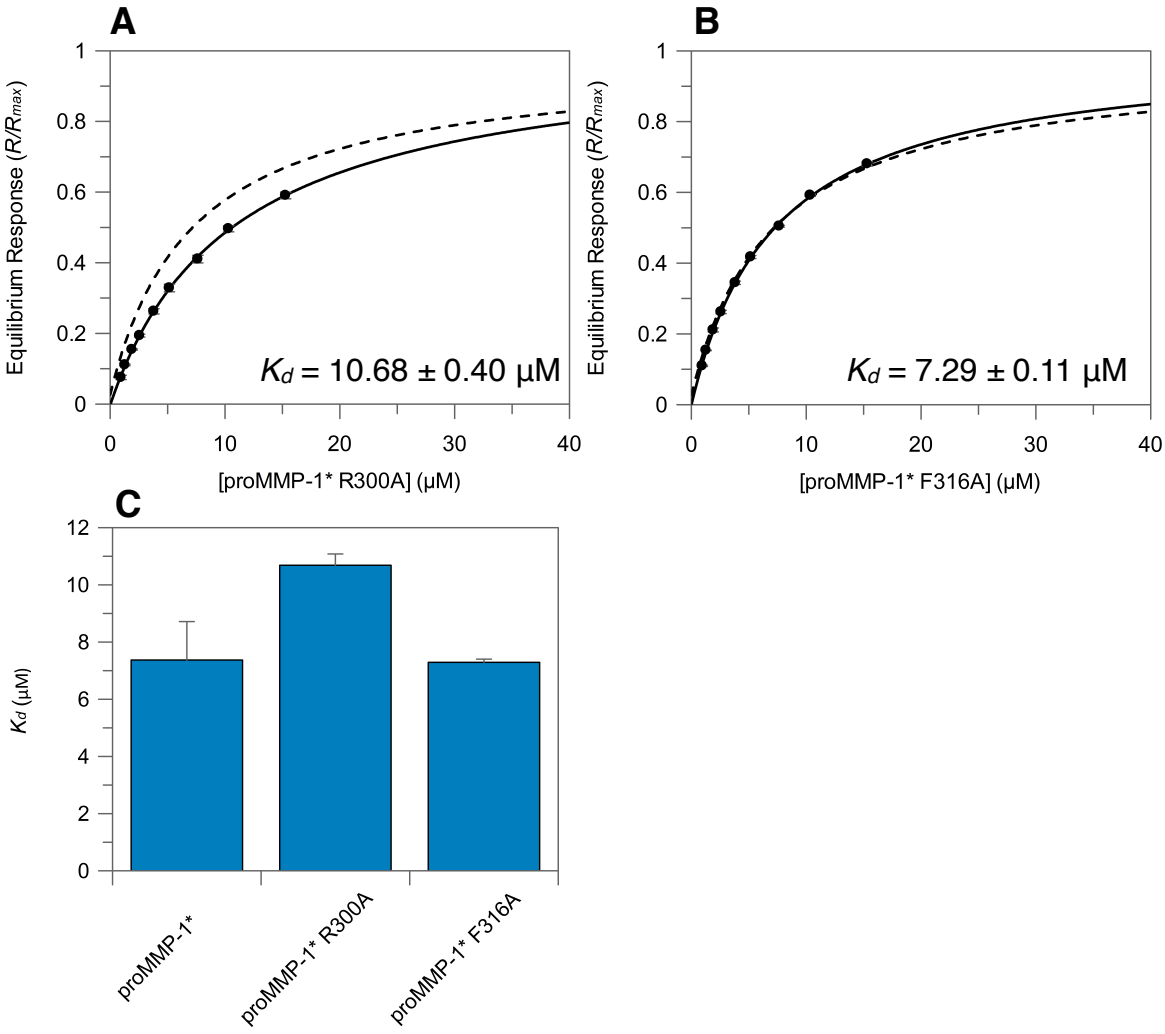



**Figure 4.8: Hyperbolic binding curves for proMMP-1* R300A and F316A from the equilibrium SPR data** (**A** and **B**). The fitted curve is of the form $R/R_{max} = c/(c + K_d)$, where c is the analyte protein concentration. The non-linear fit for proMMP-1* is shown for comparison (*dashed line*). Error bars indicate the S.D. from triplicate measurements. (**C**) Plot of equilibrium dissociation constants ($K_d$) for the binding of wild-type* and mutant proteins to immobilised THP. *P* values were determined by comparing binding affinities of mutant proteins with proMMP-1*. Statistical analysis was performed using an unpaired two-tailed Student's *t*-test (df=4) and *P* values were found to be greater than or equal to 0.05 were not considered statistically significant.

#### 4.3.4.4 Binding of Mutant MMP-1* Proteins (R300A, F301A and F316A) to Immobilised THP

MMP-1* R300A, F301A and F316A bound to α1(I)772-787 THP with apparent $K_d$ values of 2.78 ± 1.26 μM, 27.02 ± 2.98 μM and 2.69 ± 0.71 μM, respectively (Figure 4.9). Surprisingly the introduction of the highest impact HPX-1 mutation, F301A, into the full length MMP-1* form resulted in a drastic (30-fold) decrease in binding activity. Comparatively, the R300A and F316A mutations conferred conservative effects on binding (3-fold).

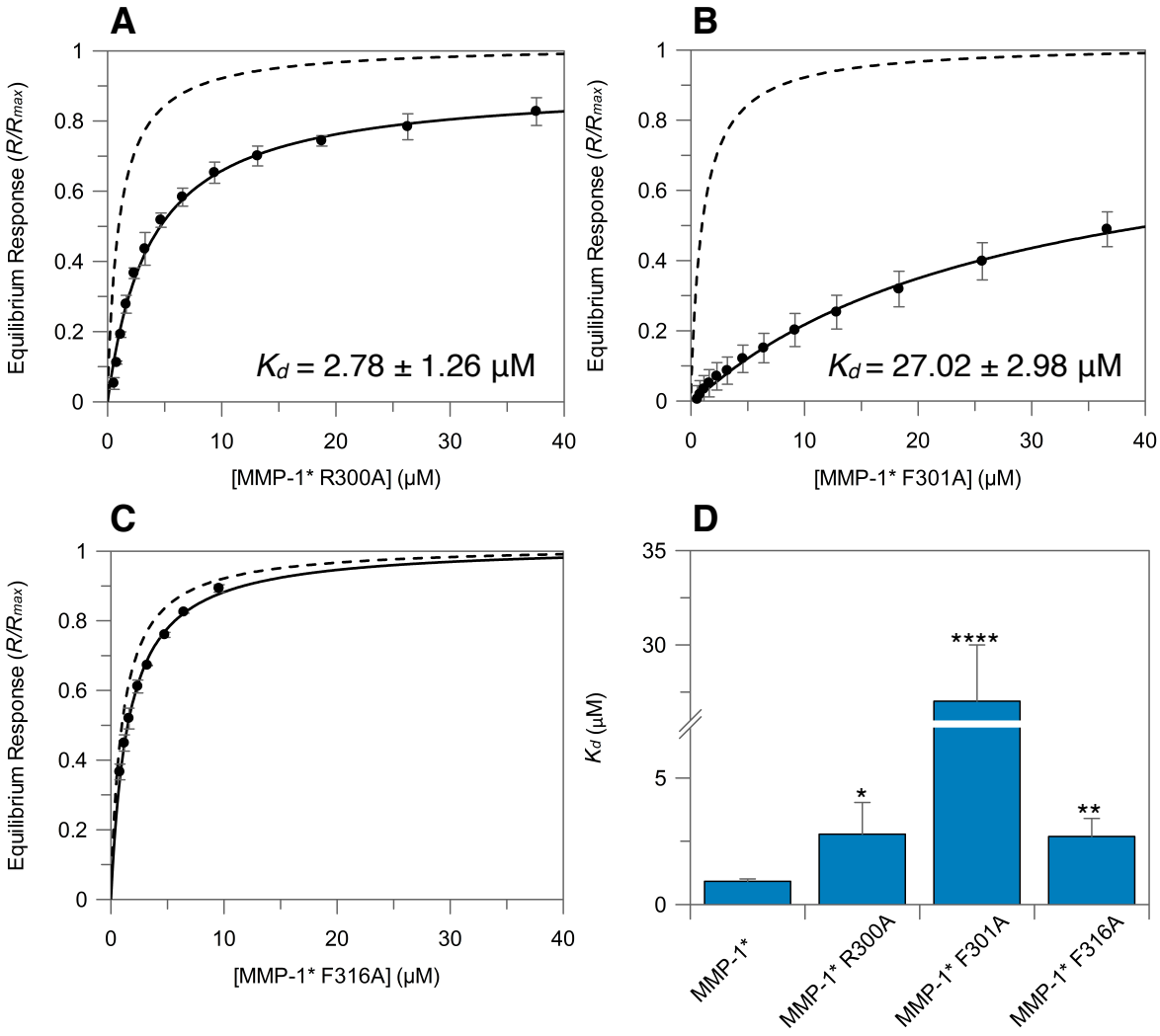



**Figure 4.9: Hyperbolic binding curves for MMP-1* R300A, F301A F316A from the equilibrium SPR data** (**A-C**). The fitted curve is of the form $R/R_{max} = c/(c + K_d)$, where c is the analyte protein concentration. The non-linear fit for MMP-1* is shown for comparison (*dashed line*). Error bars indicate the S.D. from triplicate measurements. (**D**) Plot of equilibrium dissociation constants ($K_d$) for the binding of wild-type* and mutant proteins to immobilised THP. *P*-values were determined by comparing binding affinities of mutant proteins with MMP-1*. Statistical analysis was performed using an unpaired two-tailed Student's *t*-test (df=4) and statistically significant *P*-values ($p<0.05$) are shown. The asterisk indicates statistical significance with *P*-values of $p<0.05$ (*), $p<0.01$ (**), $p<0.001$ (***) and $p<0.0001$ (****). *P*-values greater than or equal to 0.05 are not considered statistically significant.

### 4.3.5 Assay of Hydrolytic Activity using a Chromogenic Peptide

Hydrolytic activities of activated MMP-1 proteins (WT, R300A, F301A and F316A) were determined spectrophotometrically at 412 nm following the generation of 2-nitro-5-thiobenzoic acid product from the thiopeptolide substrate Ac-Pro-Leu-Gly-[2-mercapto-4-methyl-pentanoyl]-Leu-Gly-$OC_2H_5$ (Enzo® Life Sciences) (See Section 2.3.23). Active CAT-1 was used as the positive control in this study and deemed

representative of full activity (i.e. 100% relative activity). In contrast, treatment of MMP-1 recombinant proteins with 5 mM EDTA abolished hydrolytic activity, as expected, in all cases.

All 'ball' mutant MMP-1 proteins cleaved the thiopeptide substrate at a similar rate (Figure 4.10) as WT MMP-1. Specific activity values are reported in Table 4.2.

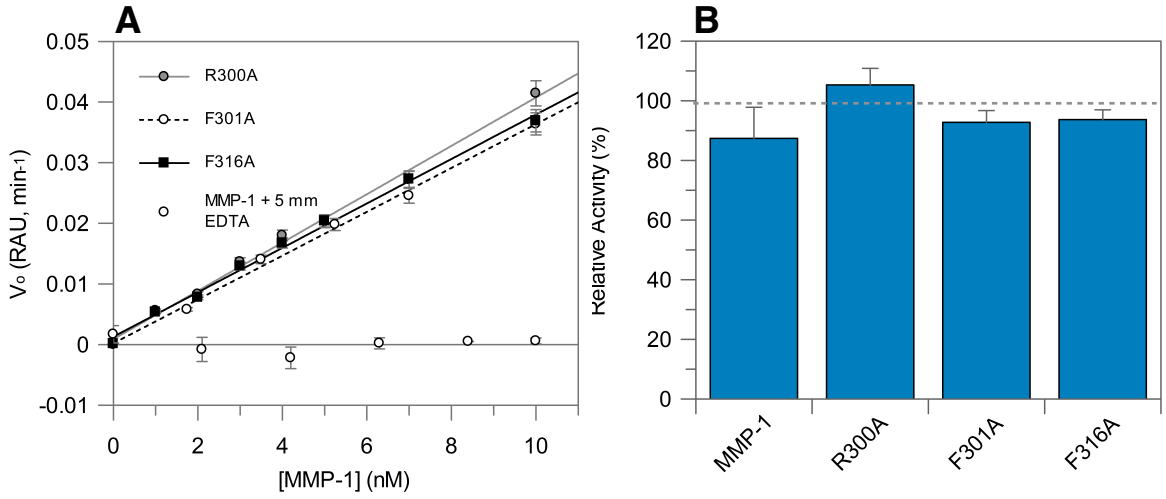



**Figure 4.10: Hydrolytic activity of MMP-1 mutants.** (**A**) Plot depicting the proportional increase of the initial reaction rates of MMP-1-catalysed thiopeptolide substrate hydrolysis by mutants. Data are mean values of three independent experiments and *error bars* indicate the SD. (**B**) Plot of relative activity (%) compared to CAT-1 (*horizontal dashed line*). *P* values were determined by comparing the hydrolytic activity of mutant proteins with CAT-1. Statistical analysis was performed using an unpaired two-tailed Student's *t*-test (df=4) and *P* values were found to be greater than or equal to 0.05, which were not considered statistically significant.

Table 4.2: Hydrolytic activity of catalytically active MMP-1 mutants determined at 37 °C

| Protein | Specific Activity (U/min/μmol)[a] |
|---|---|
| CAT-1 | 2650 ± 104 |
| WT MMP-1 | 2353 ± 190 |
| MMP-1 (R300A) | 2791 ± 140 |
| MMP-1 (F301A) | 2468 ± 99 |
| MMP-1 (F316A) | 2500 ± 81 |

[a] data from triplicate experiments is given as mean values ± SD

Interestingly, MMP-1 (R300A) showed a slight enhancement of activity when compared to both CAT-1 and MMP-1. However, it appears that specific mutations of the 'ball' joint have largely unaffected the CAT domain structure and consequently catalytic activity as would be expected since all the mutations are located in the HPX domain.

### 4.3.6 Low Resolution Structural Analysis using SAXS

The structural basis for the dramatic reduction in α1(I)772-787 THP binding by MMP-1* mutants, particularly MMP-1* (F301A), was investigated using SAXS. In addition, proMMP-1* was examined to determine whether the compact arrangement of domains in the crystal structure is representative of the conformation in solution.

#### 4.3.6.1 proMMP-1*

The apparent molecular mass of proMMP-1* was estimated as 43.8 kDa, an underestimation of 17% from the theoretical MW. This is not surprising as $I(0)$ values can not be directly determined experimentally and deviations of expected molecular mass can exceed 10% (Mylonas and Svergun 2007) The $R_g$, from Guinier analysis, was calculated at 32.8 ± 0.2 Å (Figure 4.11B). The pair distance distribution, $P(r)$, plot shown in Figure 4.11C indicates an asymmetric shape, characteristic of an elongated globular protein with a maximal particle dimension ($D_{max}$) of 90 Å. The back transformation of the *P(r)* overlaid well with the experimental data, indicating that it was a good solution for the average size and shape of proMMP-1* in solution. CRYSOL (Svergun, Barberato et al. 1995) was used to compare theoretical scattering of proMMP-1* in the crystal structure conformation with the experimental scattering data of the proMMP-1* solution conformation (Figure 4.11D). At low angles the curves overlay and show a good agreement between both structures, however the curves within the medium $q$ range (> 0.2 $Å^{-1}$) show variation and may indicate changes in relative domain orientation between the two structures.

A representative low resolution *ab initio* DAMMIF (Franke and Svergun 2009) model of the proMMP-1* conformation is shown in Figure 4.14. SUPCOMB (Kozin and Svergun 2001) was used to dock the high resolution crystal structures within the molecular envelope. This shows that the model largely agrees with the compact arrangement of the domains in the crystal structure, as shown by the reasonably low normalised spatial discrepancy (NSD) value of 1.65. The discrepancies that do exist between the two may be due to the fact that a significant proportion of the electron

density map for proMMP-1* is missing from the crystal structure, namely that for the N-terminal residues (F20-V31), the bait region (N46-S57), and the PRO-CAT linker (Q99-N106) (Jozic, Bourenkov et al. 2005).

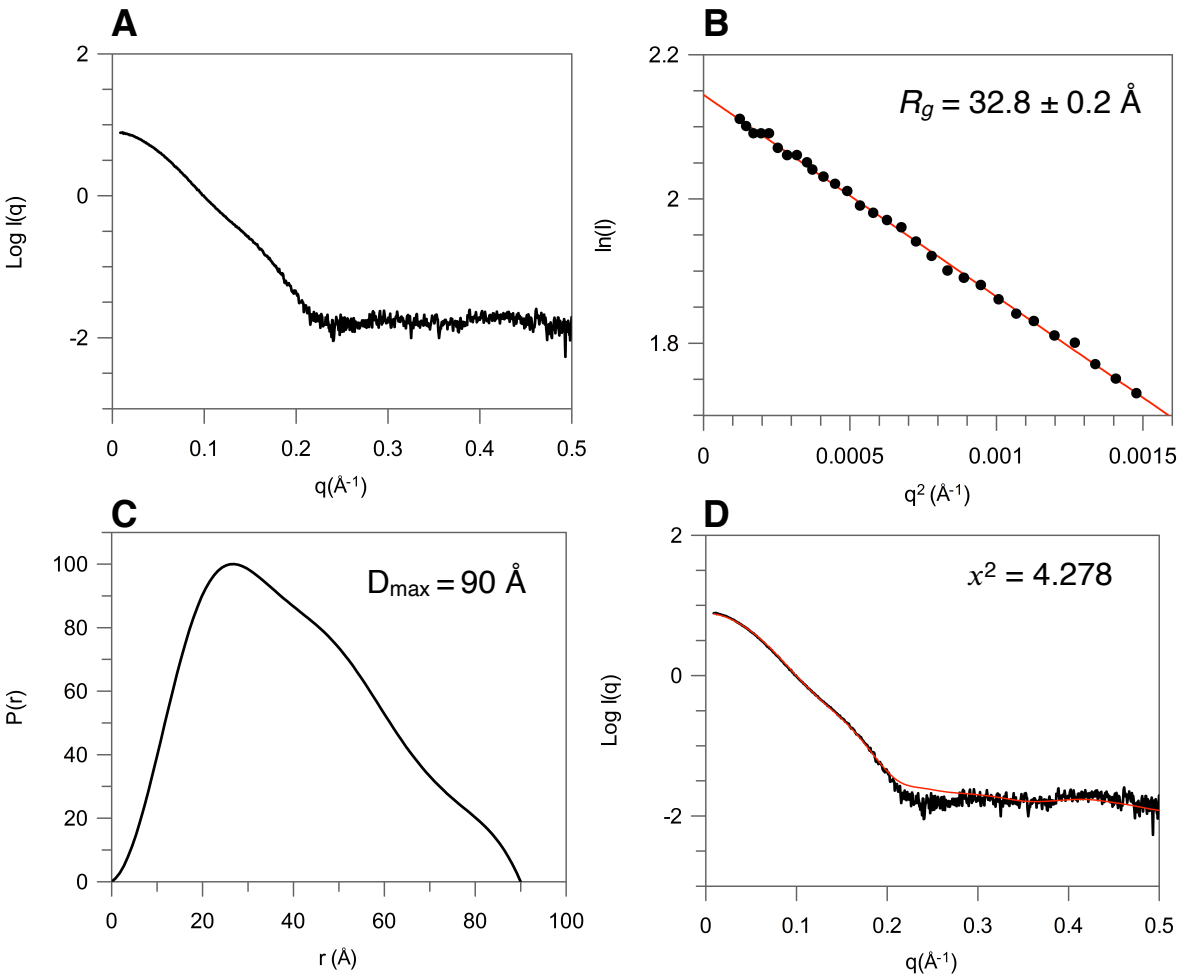



**Figure 4.11: SAXS analysis of proMMP-1*.** (**A**) Experimental SAXS scattering profile of zymogen proMMP-1* after solvent subtraction. (**B**) Guinier plot of the low-angle region limited by $qR_g<1.3$. (**C**) P(r) function generated by GNOM (Svergun 1992). (**D**) Theoretical scattering curve (*red*) calculated from the crystal structure (PDB accession code: 1SU3) overlaid on the experimental scattering data (*black*).

### 4.3.6.2 MMP-1*

The molecular mass of MMP-1* was estimated as described in Section 4.3.5.1 at 34.8 kDa. This is an underestimation of 18% but clearly indicates a reduction in mass of 8.9 kDa, which is accounted for by the loss of the ~ 9-kDa PRO domain upon activation. Furthermore, zymogen activation is concomitant with a significant reduction in $R_g$ of 12% to 28.7 ± 0.2 Å when compared with proMMP-1* (Figure 4.12A), which agrees with the previous report of 29 ± 1 Å for MMP-1* (Bertini, Fragai et al. 2009). The

absence of the PRO domain also results in a small decrease in the $D_{max}$ from 90 to 88 Å (Figure 4.12C).

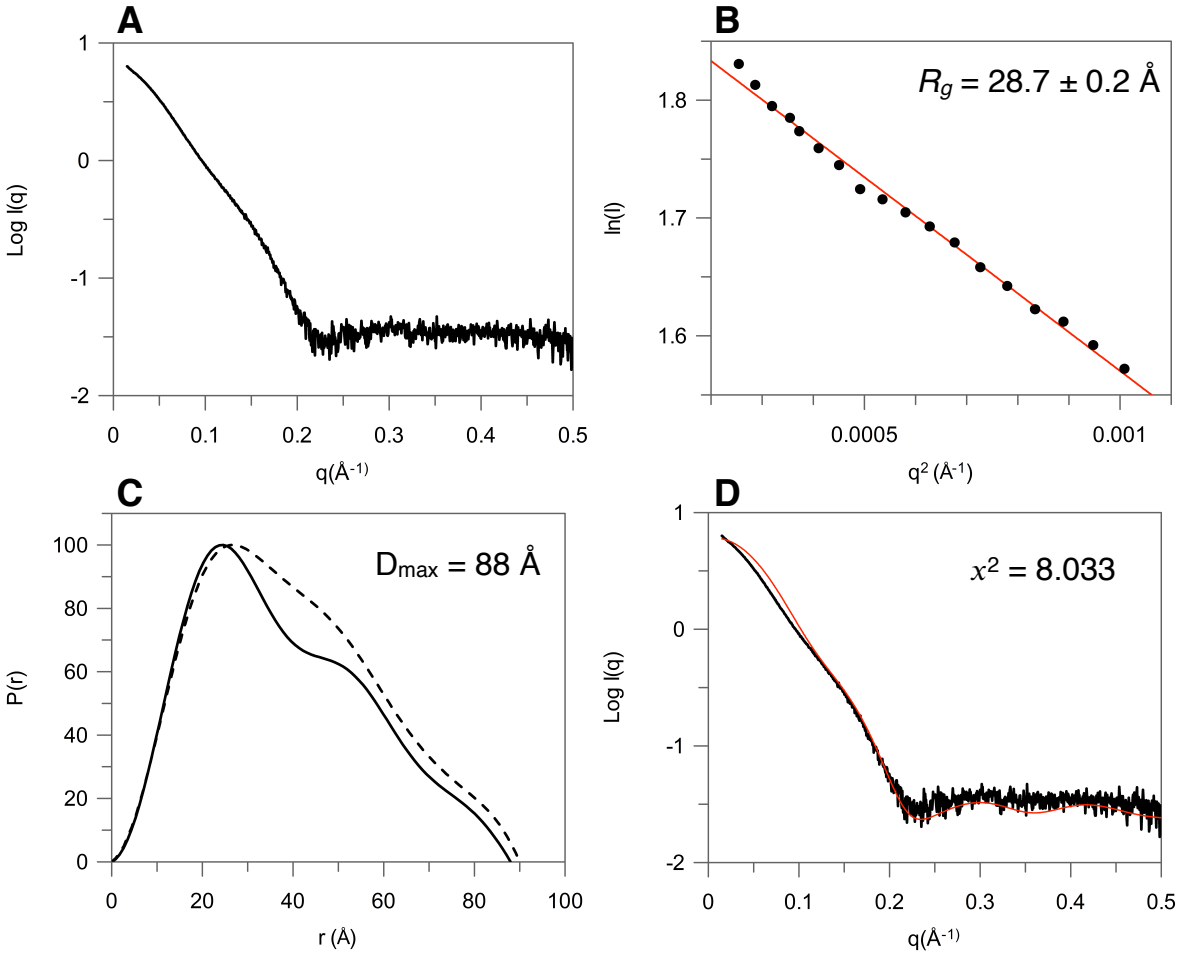



**Figure 4.12: SAXS analysis of MMP-1*.** (**A**) Experimental SAXS scattering profile of mature MMP-1* after solvent subtraction. (**B**) Guinier plot of the low-angle region limited by $qR_g<1.3$. (**C**) P(r) function generated by GNOM (Svergun 1992). The proMMP-1* P(r) distribution plot is shown for comparison (*dashed line*). (**D**) Theoretical scattering curve (*red*) calculated from the crystal structure (PDB accession code: 2CLT) overlaid on the experimental scattering data (*black*).

The asymmetric *P(r)* plot indicates significant variation when compared to that of proMMP-1*. The distinctive shoulder at ~55 Å is indicative of a discrete domain structure and implies a slight separation of CAT and HPX domains as observed by Bertini and co-workers (Bertini, Fragai et al. 2009). Variations in the CRYSOL (Svergun, Barberato et al. 1995) output throughout the sampled *q* range are also apparent with significant differences in the low *q* range (< 0.2 $Å^{-1}$). This is consistent with an increased experimental $R_g$ value of 28.6 Å when compared with the theoretical $R_g$ value of 25.95 Å for the crystal structure (Bertini, Fragai et al. 2009), and is further highlighted by the DAMMIF (Franke and Svergun 2009) generated molecular envelope,

which is considerably larger than expected from its compact crystal structure (Figure 4.14). Superimposition of these low and high resolution structures results in a considerably higher NSD value of 2.07 compared to proMMP-1*, and leaves an entire lobe of the solution-state model unaccounted for.

Control models were generated (Figure 4.14) using artificial scattering data synthesised from the crystal structure coordinates of proMMP-1* (PDB accession code: 1SU3) and MMP-1* (PDB accession code: 2CLT) to ensure that the molecular envelope discrepancy observed for MMP-1* was not an artefact of the modelling process. These control models were in close agreement with the crystal structure coordinates with NSD values of 1.24 and 1.19, respectively.

#### 4.3.6.3 MMP-1* Mutants R300A, F301A and F316A

SAXS analysis of MMP-1* mutants R300A, F301A and F316A indicated $R_g$ values of 34.7 ± 0.03 Å, 33.6 ± 0.04 Å and 30.9 ± 0.01 Å, respectively (Figure 4.13). The greatest increases observed were 21% for MMP-1* (R300A), and 17% for MMP-1* (F301A) with comparable $R_g$ values to proMMP-1 despite the absence of the PRO domain in these mutants. Similarly, increases over the WT value of 88 Å were observed with $D_{max}$ values of 104 Å, 90 Å and 93 Å for the R300A, F301A and F316A mutants, respectively. Furthermore, the *P(r)* plots of these mutants indicate that the distinct shoulder observed for MMP-1* becomes much more exaggerated as shown by the increasing distribution of larger distances. This implicates a significant separation of the CAT and HPX domains. In light of the adversely impacted THP binding seen by MMP-1* F301A, *ab initio* modelling was conducted for this mutant. The low resolution envelope exhibits two discrete lobes (Figure 4.14), which may be attributed to permanent dislocation of the CAT and HPX domains brought about by mutation of residues at the CAT:HPX interface.

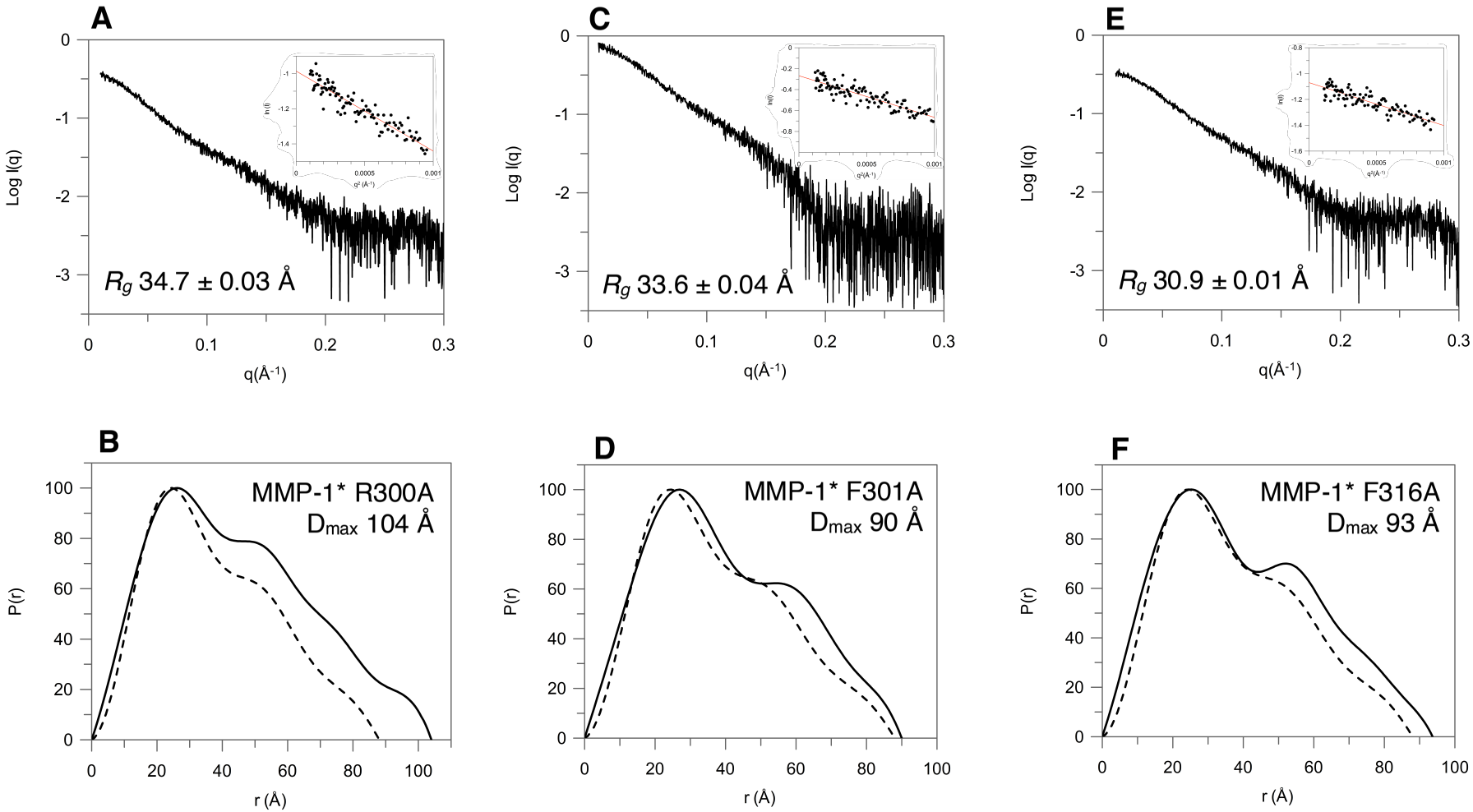



**Figure 4.13: SAXS analysis of MMP-1* mutants.** (**A**, **C** and **E**) Experimental SAXS scattering profile of mature MMP-1* mutants after solvent subtraction. Guinier plot of the low-angle region limited by $qR_g$<1.3 (*Inset*). (**B**, **D** and **F**) P(r) function generated by GNOM (Svergun 1992). The MMP-1* P(r) distribution plot is shown for comparison (*dashed line*).

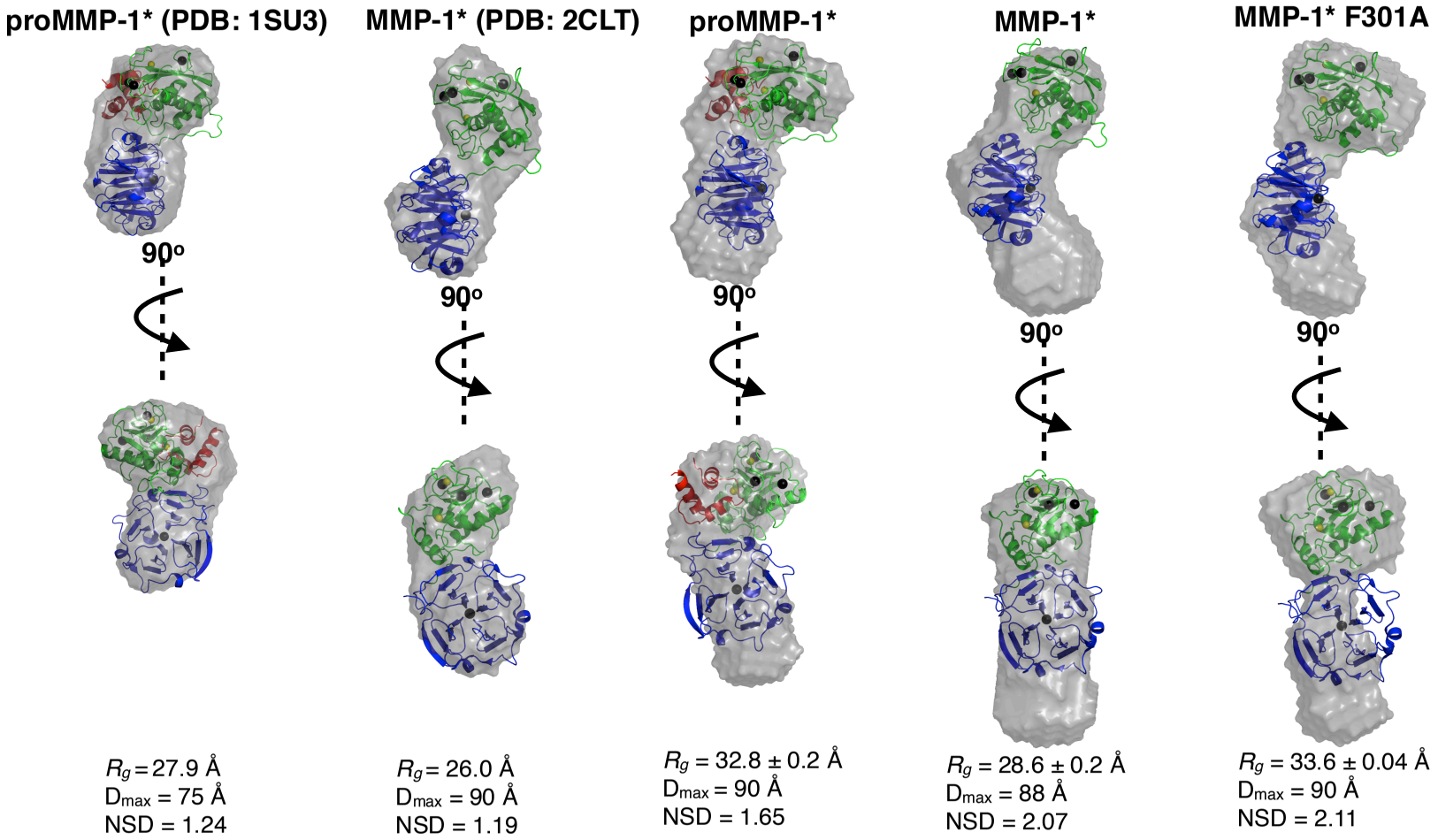



**Figure 4.14: Orthogonal views of the average low resolution molecular envelopes.** *Ab initio* modelling (*grey*) from experimental and theoretical scattering data with crystal structures overlaid using the program SUPCOMB (Kozin and Svergun 2001). For MMP-1* and MMP-1* F301A a manual overlay was performed over the crystal structure of MMP-1* as automated overlays gave unacceptably poor fits (NSD values of 2.07 and 2.11, respectively). Note that for these proteins unique orientations of the MMP-1* crystal structure with respect to the *ab initio* low resolution structures could not be identified due to the similar sizes of the CAT and HPX domains. Thus, their relative positioning here is arbitrary and only serves to highlight the discrepancy between solution-state and solid-state observations. In the crystal structures, the PRO domain is coloured *red*, the CAT domain in *green* and the HPX domain in *blue*.

## 4.4 Conclusions

Mutant pET3a:proMMP-1* and proMMP-1 constructs were successfully produced by site-directed mutagenesis and transformed into *E. coli* BL21-CodonPlus® (DE3)-RIPL cells, where the recombinant proteins were expressed at comparable levels to that of the WT enzyme. The standard refolding and purification strategy for proMMP-1 production, devised previously (Chapter 2), resulted in sufficient material for the assays described herein. Successful activation of proMMP-1 with acquisition of full hydrolytic activity against a peptolide substrate, when compared to CAT-1, was taken as suitable evidence that each protein was folded correctly and suitable for further study.

Characterisation of homotrimeric collagen THP binding by full length proMMP-1* and MMP-1* enzymes was completed using an SPR-based assay. Interestingly, binding activity was detected for the proMMP-1* zymogen. Previous studies, using gel filtration chromatography and ELISA, have concluded that proMMP-1 is unable to bind collagen substrates, an activity acquired upon PRO domain removal (Welgus, Jeffrey et al. 1985, Murphy, Allan et al. 1992). However, our experimental observations result from a particularly sensitive technique likely to detect this interaction, which we report as significantly weaker (~10-fold) than that observed for mature MMP-1*. Furthermore, dissociation constants previously determined for MMP-1* bound to both homo- and heterotrimeric mouse tail tendon type I collagen display similar values to those observed here, 1.3 ± 0.2 and 1.3 ± 0.3 μM, respectively (Han, Makareeva et al. 2010). Increases in $K_d$ for MMP-1* were observed for both mutants R300A and F316A. However, the highest impact HPX-1 mutation, F301A, had significant adverse effects on THP binding (See Chapter 3), decreasing MMP-1* affinity 30-fold. This suggests that this F301A is an important contributor to collagen recognition despite being located within the CAT:HPX interface, a position that would appear to prevent involvement in collagen binding. The seemingly moderate detrimental effects seen for R300A and F316A may therefore result from distortions of the peptide backbone affecting positioning of the F301A sidechain, rather than direct interactions between these residues and the substrate, or more likely via disruption of the cooperative domain binding mechanism.

These mutations also appear to have affected the structure of MMP-1*, with increases in hydrodynamic radii detectable by analytical gel filtration. Therefore, the structural basis for the reduction in binding activity compared with MMP-1* was investigated using SAXS. Initial analysis indicated that both proMMP-1* and MMP-1*

are elongated proteins with solution structures differing from that of the compact crystal structure conformations, this is particularly evident for MMP-1* (Figure 4.14). Most interestingly, *ab initio* modelling produced a considerably larger molecular envelope than expected from the crystal structure with additional regions of density unaccounted for when superimposed with the high resolution structure. Thus, the elongated low-resolution solution structure of MMP-1* is indicative of an equilibrium between the compact state seen in crystallographic analyses and a more extended state. This is in agreement with the previous study by that used NMR and SAXS to demonstrate that the compact arrangement of domains for MMP-1* is not fully representative of the conformation in solution (Bertini, Fragai et al. 2009). Instead, for approximately one third of the time, the enzyme exists in an elongated state with an extended CAT and HPX domain arrangement. This investigation shows the apparent dislocation of the CAT and HPX domains in mature MMP-1* is further exacerbated by mutation of residues that constitute the 'ball' joint at the CAT:HPX interface. The *P(r)* distance distribution function and $R_g$ values for R300A, F301A and F316A MMP-1* mutants are indicative of an extended structural state, which may have disrupted the dynamic interdomain interaction, that is a feature of the WT enzyme. This is more evident in the low resolution model of MMP-1* (F301A), which exhibited a dislocated state with two discrete lobes and a loss of density that correlates with a compact state. Furthermore, it has been demonstrated that the compact conformation of MMP-1* is biologically relevant, as complete disruption of the CAT:HPX interface reduces collagen binding activity. Thus, MMP-1 collagenase activity most probably depends on a subtle equilibrium between these compact and dislocated states, the second of which exposes a collagen binding exosite involving F301A.

# CHAPTER 5: ROLE OF THE DOMAIN-DISLOCATED STATE IN PROTEOLYTIC SUSCEPTIBILITY AND COLLAGEN BINDING

## 5.1 Introduction

In the previous chapter the data suggests that MMP-1* undergoes apparent transient separation of domains. This observation is supported by previous reports of MMP interdomain flexibility, where domain reorientation may be mediated by the variable-length flexible linker (Tsukada & Pourmotabbed, 2002). The findings in Chapter 4 also implicate contributions from residues at the CAT-HPX interface in domain separation impacting substrate recognition. Restriction of domain mobility, by tethering the CAT and HPX domains, could be achieved by introduction of an interdomain disulfide bond. If successful, any disruption to the conformational equilibrium could be assessed.

Disulfide bond formation generally occurs in the ER by oxidation, consequently disulfide bonds are mostly found in extracellular, secreted and periplasmic proteins where they contribute to protein mechanical stability and function. This stabilising property makes non-native disulfide-bonding a powerful strategy for engineering additional conformational stability using site-directed mutagenesis. Disulfide-bond engineering has successfully been used to investigate functional aspects of structural heterogeneity, this includes receptor activation by the chemokine Lymphotactin (Tuinstra, Peterson, Elgin, Pelzek, & Volkman, 2007), σ factor binding by the β' subunit of *E. coli* RNA polymerase (Anthony, Dombkowski, & Burgess, 2002) and ligand binding by the I domain of integrin αLβ2 (Shimaoka et al., 2001).

Examination of the MMP-1* structure indicates a position at the CAT-HPX interface, which is of a suitable distance (Figure 5.1A *inset*) to allow engineering of such a novel disulfide bond (Figure 5.1B). The Cα atoms of S243 of the CAT domain and S318 of the HPX domain are separated by 5.2 Å, with correct side chain orientation this is an optimal distance for disulfide bond formation as assessed by the Disulfide by Design 2 program (Dombkowski, 2003). Furthermore, this program predicts the geometry of cysteine pairs in a favourable orientation for disulfide bond formation (Figure 5.1C), this is consistent with the proximity and geometry constraints for disulfide formation, defined by Petersen (1999).

The study described in this chapter uses site-directed mutagenesis and assay to investigate the role of the dislocated state by tethering the CAT and HPX domains at the interface. Limited proteolysis was used to probe domain structure and assess successful enzyme maturation. SPR was used to assess the impact of domain tethering on substrate recognition using a synthetic THP substrate, and a chromogenic assay was used to characterise and quantitate the hydrolytic activity of the domain-stapled mutant, S243C, S318C. SAXS was used to investigate domain arrangement and define global structural parameters for comparison to both solution and solid-state structures. Finally, preliminary crystallisation attempts are described.

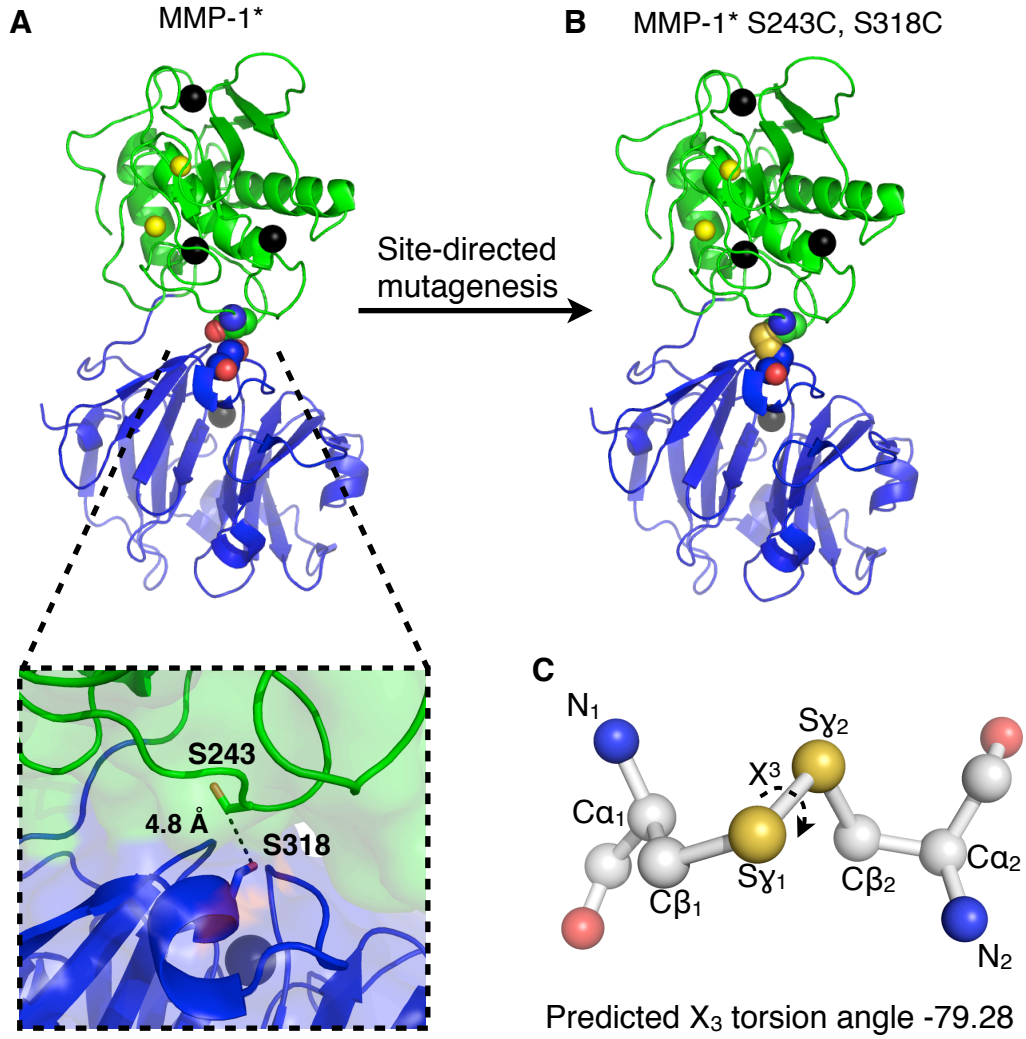



**Figure 5.1: Domain stapling of MMP-1* using site-directed mutagenesis.** The crystal structure of MMP-1* (PDB accession code 2CLT) is shown as a ribbon diagram. Zinc atoms are depicted as *yellow spheres* and calcium ions are coloured *black*. (**A**) Serine residues targeted for mutation are shown as *spheres* and are coloured according to atom. Internuclear distance between S243 and S318 γ-oxygen atoms is shown *inset*. (**B**) Modelled cysteine substitution at positions 243 and 318, shown as *spheres* and coloured according to atom. (**C**) Schematic ball and stick representation of the non-native disulfide bond designed using the Disulfide by Design 2 program v 2.06 (Dombkowski, 2003). The $X_3$ torsion angle is defined by the $C_\beta$-$S_\gamma$-$S_\gamma$-$C_\beta$ bonds, and is consistent with the $X_3$ bond angles of disulfides, which show peaks in distribution at +100 and -80 (Petersen et al., 1999). Atoms of the CAT domain and HPX domain are denoted (1) and (2), respectively. Rendered using the Pymol Molecular Graphics System (version 1.5.0.4.), Schrödinger, LLC.

## 5.2. Materials and Methods

The majority of methods necessary to investigate domain-stapling effects on MMP-1 structure and function have previously been described in Chapters 2-4. Oligonucleotide design and preparation was undertaken as described in Section 2.3.4. Site-directed mutagenesis of target residues was completed as described in Section 2.3.5. THP substrate-binding was assayed and analysed using SPR as described in Section 3.2.2.2 and 3.2.2.3. MMP-1* S243C, S318C global shape parameter analysis was completed using SAXS as described in Section 4.2.3.

### 5.2.1 Limited proteolysis using CAT-1

Purified mature WT MMP-1 samples (40 nM) were incubated for 24 hours in a 37°C waterbath, prior to application at 1 mL/min to a pre-equilibrated 5 mL HisTrap FF column pre-packed with Ni Sepharose 6 Fast Flow resin using the ÄKTAprime system. Unbound CAT-1 was collected in the flowthrough and concentrated, using centrifugal filter units with an appropriate MWCO (Millipore) using a Beckman Coulter Allegra 25R refrigerated centrifuge (rotor AT-14-10) pre-chilled to 4 °C. Samples were resuspended after each spin (3000 rpm for 1-10 minutes). Bound proteinaceous material was eluted with a 0-100% gradient of TNC buffer supplemented with 500 mM imidazole and collected in 1 mL fractions for further SDS-PAGE analysis (Section 2.3.19).

A three-fold molar excess of CAT-1 was added to proMMP-1* proteins in TNC Buffer. Samples were then incubated at 37 °C for 0-20 hours. Reactions were stopped with 2X SDS-PAGE reducing loading buffer as described in Section 2.3.20.2.

### 5.2.2 Crystallisation of MMP-1* (S243C, S318C)

Crystallisation of MMP-1* (S243C, S318C) at a concentration of 120 µM in TNC buffer was performed in an attempt to obtain atomic resolution of the domain-stapled species. The following sparse matrix crystallisation screens (Molecular Dimensions) were used: Morpheus, JCSG-plus and PACT-premier. Sitting drops were set up in a 96 condition MRC-style two well plate using the Cartesian dispensing system (Genomic Solutions) run with Honey Bee software (Digilabs). The robot dispensed 200 nL of sample and 200 nL of the well solution. The trays were stored at 16°C and checked at regular intervals using polarizing light microscopy.

## 5.3. Results

### 5.3.1 Mutagenesis

Site-directed mutagenesis of the pET-3a proMMP-1* plasmid construct was completed using the Quikchange II XL Site-directed Mutagenesis kit (Agilent Technologies). Mutagenic efficiency levels were determined as previously described and were consistent with previous observations (Section 4.3.1). DNA sequence analysis confirmed successful incorporation of mutagenic bases and hence single cysteine amino acid substitutions at positions 243 and 318 in proMMP-1* (See Appendix 12). Once the proMMP-1* S243C, S318C mutations were confirmed, a A219E revertant was then generated in the same manner to restore hydrolytic activity.

### 5.3.2 Protein Production

Catalytically impaired proMMP-1* S243C, S318C and the equivalent revertant constructs were transformed into *E. coli* expression strain BL21-CodonPlus (DE3)-RIPL and screened for IPTG-induced recombinant protein expression. Proteins of ~52 kDa were expressed, which is consistent with the expected mass.

As shown previously (Chapters 2, 3 and 4) all target proteins were found to deposit into bacterial inclusion bodies and were therefore solubilised, refolded and purified in accordance with standard procedures (See Chapter 2). Despite refolding under the same conditions as the wild-type zymogen, precipitated protein levels were found to increase impacting negatively on the yield of purified protein obtained from 500 mL of culture. This was anticipated due to the increased potential of incorrect inter- and intramolecular disulphide bond formation. However, sufficient milligram quantities were obtained by pooling purified protein grown from multiple cultures, and analytical gel filtration chromatography and non-reducing SDS-PAGE was used to confirm that no soluble disulphide-bonded multimers were present (Figure 5.2).

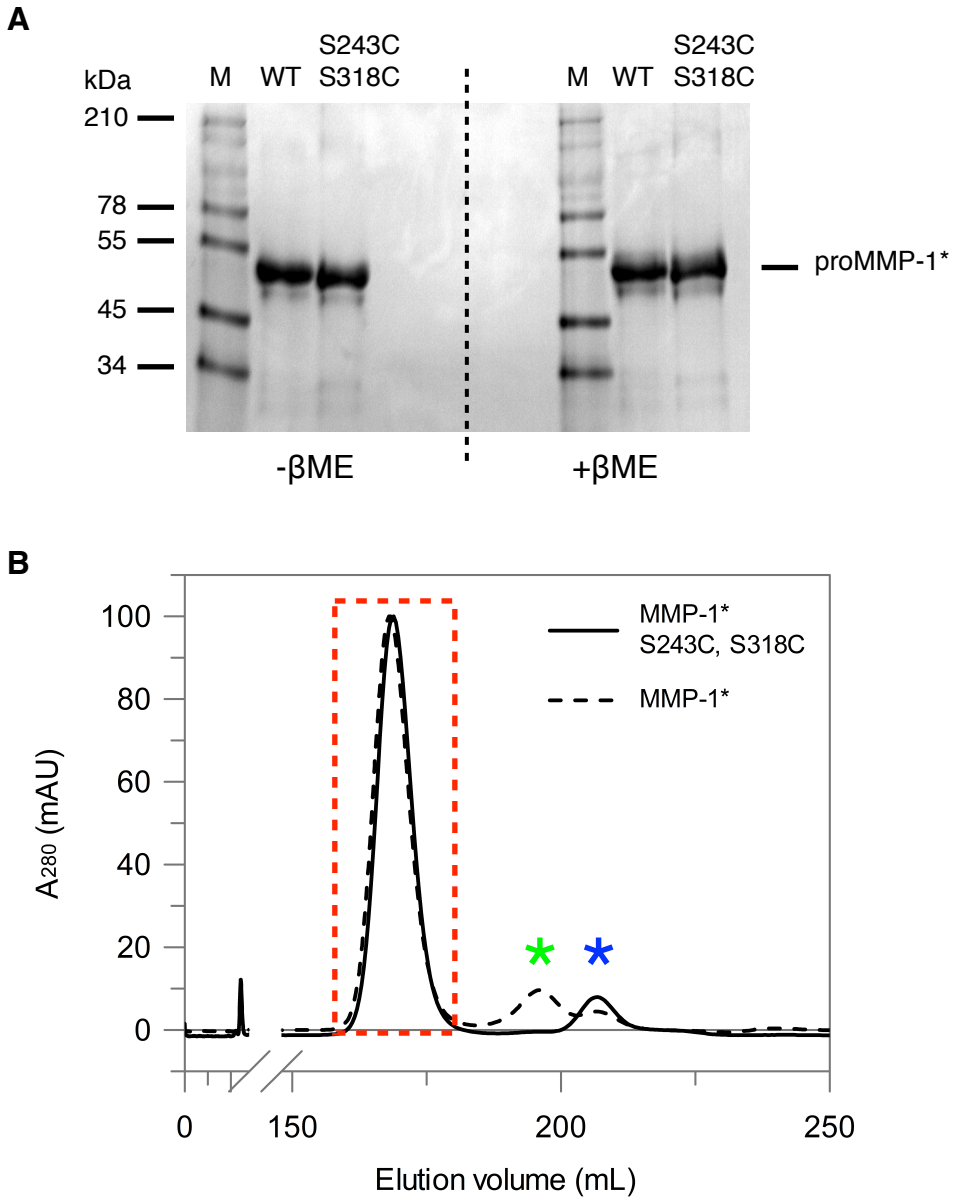



**Figure 5.2: Determining the presence of disulphide-bonded multimers.** (**A**) Lack of disulphide-bonded multimers in the S243C, S318C proMMP-1* mutant as shown by oxidized and reduced SDS-PAGE and visualised on a 16.5% (w/v) Tris-Tricine SDS-PAGE gel. Molecular weight was estimated using SeeBlue Plus2 pre-stained protein ladder (**M**). (**B**) Elution profile of MMP-1* S243C, S318C by analytical gel filtration using a HiLoad Superdex 75 16/60 prep grade column. Purified mature protein is shown with a *dashed red box*. The elution profile of MMP-1* is shown for comparison purposes (*dashed line*). Elution position of PRO-1 and CAT-3 is indicated with a *blue* (*) and *green asterisk* (*), respectively.

### 5.3.3 Low resolution structural analysis using SAXS

SAXS was used to determine whether the solution state of the proposed domain-stapled mutant possessed a compact arrangement of domains similar to that of the crystal structure. The $R_g$, from Guinier analysis, was calculated at 27.8 ± 0.2 Å (Figure 5.3A), which is 3% lower than that observed for the experimentally-derived MMP-1* solution structure but 7% greater than that estimated by Bertini and co-workers from the crystal structure (Bertini et al., 2009). Analysis of the P(*r*) plot (Figure 5.3B) shows the asymmetric shape that appears to be characteristic of MMP-1 (Chapter 4) with a shoulder at ~55 Å indicative of bilobal protein structure. However, a reduction in $D_{max}$ of 6 Å is observable, which with the $R_g$ value, indicates a general reduction in overall dimensions. Furthermore, the *ab initio* molecular envelope generated using DAMMIF (Franke & Svergun, 2009) and superimposed with the crystal structure gave a reasonably low NSD value of 1.73. This is a better fit than that observed for the MMP-1* solution data (NSD 2.07) and is suggestive of a more compact two lobe state.

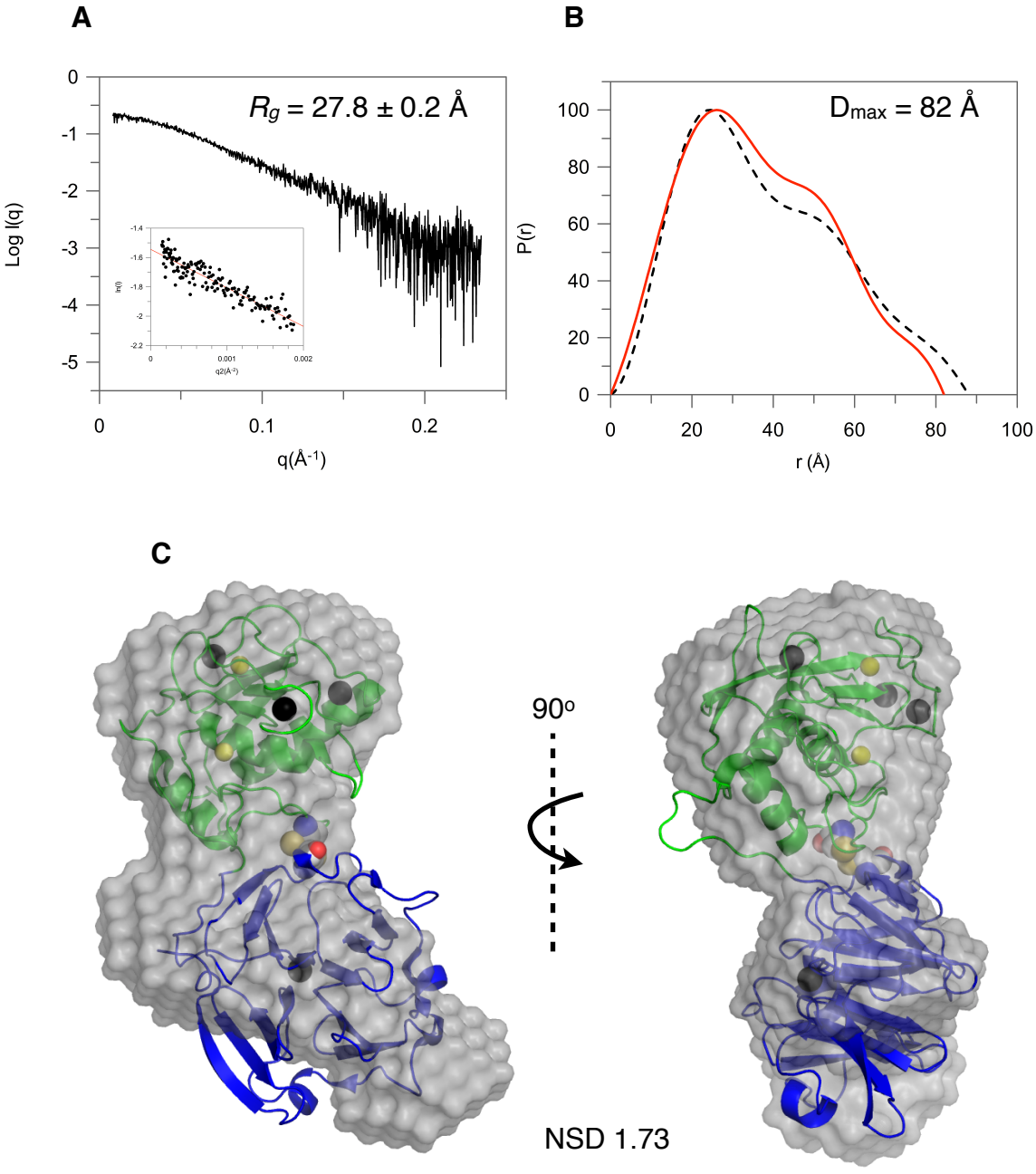



**Figure 5.3: SAXS analysis of MMP-1* S243C, S318C.** (**A**) Experimental SAXS scattering profile of the domain-stapled mature MMP-1* mutant after solvent subtraction. Guinier plot of the low-angle region limited by $qR_g$<1.3 (*Inset*). (**B**) P(r) function generated by GNOM (Svergun, 1992). The domain-stapled mutant is shown in *red* and MMP-1* (*dashed black line*) is shown for comparison (Chapter 4). (**C**) *Ab initio* modelling from experimental scattering data with orthogonal views of the MMP-1* crystal structure (PDB accession code: 2CLT) overlaid using the program SUPCOMB (Kozin & Svergun, 2001). In the crystal structure, the CAT domain is coloured in *green* and the HPX domain in *blue*. The cysteine mutations (S243C, S318C) at the CAT-HPX interface are shown as *spheres* and coloured according to atom.

### 5.3.4 Limited proteolysis of proMMP-1* S243C, S318C by CAT-3 and CAT-1

The domain-stapled mutant was incubated with an equimolar concentration of CAT-3 to determine accessibility and cleavage of the Gln80-Phe81 bond in the PRO-CAT linker, concomitant with PRO-domain loss. Comparison with both the catalytically-impaired zymogen and domain-dislocated proMMP-1* 'ball' mutant F301A (See Chapter 4), indicated that proMMP-1* S243C, S318C was resistant to specific proteolysis and could be activated in the same way (Figure 5.4A).

The linker of active MMP-1 is susceptible to autolysis producing two discreet products, CAT-1 and HPX-1 (Li et al., 1995). The fortuitous observation that HPX-1 bound immobilised nickel (Chapter 2) provided a method enabling CAT-1 isolation, which was then used to probe CAT-HPX linker accessibility of zymogen forms (Figure 5.4B). After 20 hours incubation both proMMP-1* and the F301A mutant showed degradation characteristic of linker cleavage. In contrast, there was a distinct absence of both proCAT-1* and HPX-1 products when proMMP-1* S243C, S318C was incubated with catalytically active CAT-1. This suggests that the domain-stapled mutant is significantly more resistant to hydrolysis of the CAT-HPX interdomain linker.

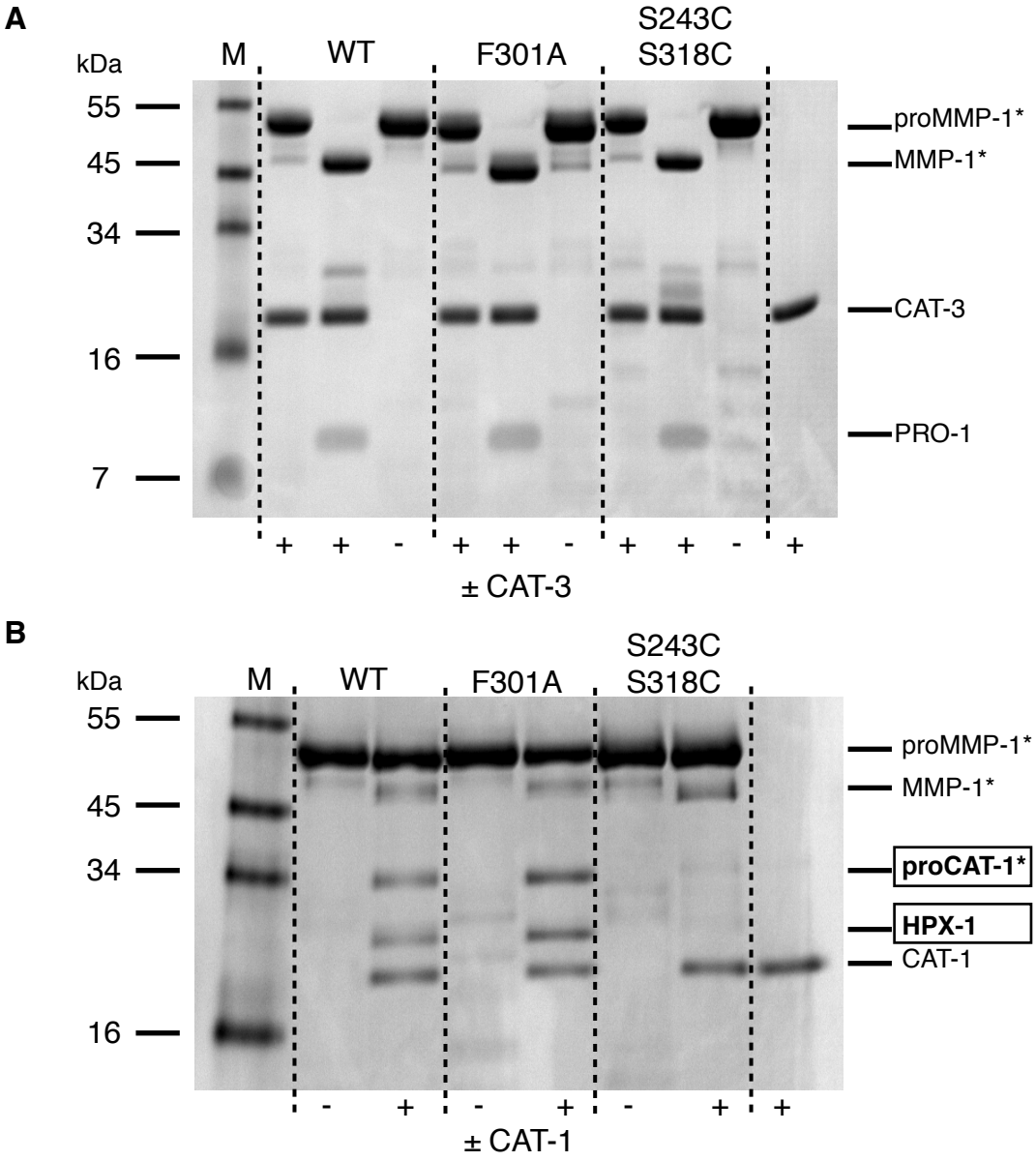



**Figure 5.4: Limited proteolysis of recombinant proMMP-1* proteins by CAT-3 and CAT-1.** (**A**) Purified proMMP-1* (**WT**) and mutant proteins were incubated with CAT-3 in TNC buffer at an equimolar concentration for 0 and 20 hours at 37°C. (**B**) Purified proMMP-1* (**WT**) and mutant proteins were incubated with CAT-1 in TNC buffer at a 3:1 molar concentration for 0 and 20 hours at 37°C. The products indicative of linker cleavage are indicated by *boxed titles*. All reactions were visualised on 16.5% (w/v) Tris-Tricine SDS-PAGE gels. Molecular weight was estimated using SeeBlue Plus2 pre-stained protein ladder (**M**).

### 5.3.5 Functional analysis by SPR

Equilibrium SPR analysis was used to characterise the binding affinity of both the proMMP-1* zymogen and mature MMP-1* domain-stapled proteins for immobilised α1(I)772-787 THP. Triplicate injections were performed using a suitable range of analyte concentrations (0.2 μM-20 μM) and affinities were determined as described in Section 3.2.2.2 and 3.2.2.3.

The binding of the S243C, S318C zymogen form was significantly impaired prohibiting the calculation of a $K_d$ value. Upon PRO-domain removal, the domain-stapled mutant bound to α1(I)772-787 THP with an apparent $K_d$ value of 8.9 ± 0.6 μM (Figure 5.5A). Cysteine substitution at positions 243 and 318 clearly diminished THP binding, with a ~10-fold reduction in $K_d$ when compared to MMP-1*. However, the impact on collagen recognition was not as severe as that observed for the F301A domain-dislocation mutant but more so than for R300A and F316A (Figure 5.5B).

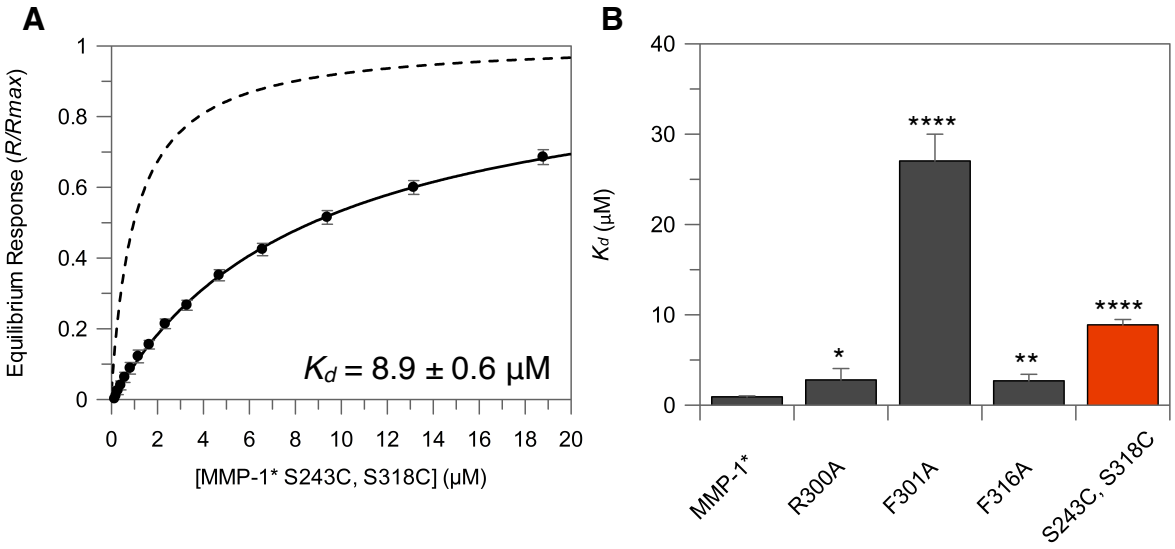



**Figure 5.5: Equilibrium binding of MMP-1* S243C, S318C to α1(I)772-787 THP.** (**A**) Hyperbolic binding curves for MMP-1* from the equilibrium SPR data. The fitted curve is of the form $R/Rmax = c/(c + K_d)$, where $c$ is the analyte protein concentration. The non-linear fit for MMP-1* is shown for comparison (*dashed line)*. Error bars indicate the S.D. from triplicate measurements. (**B**) Plot of equilibrium dissociation constants ($K_d$) for the binding of MMP-1* and mutant proteins to immobilised THP. *P*-values were determined by comparing binding affinities of mutant proteins with MMP-1*. Statistical analysis was performed using an unpaired two-tailed Student *t*-test (df = 4) and statistically significant *P*-values ($p<0.05$) are shown. The asterisk indicates statistical significance with *P*-values of $p<0.05$ (*), $p<0.01$ (**), $p<0.001$ (***) and $p<0.0001$ (****). *P*-values greater than or equal to 0.05 are not considered statistically significant.

### 5.3.6 Assay of hydrolytic activity using a chromogenic peptide substrate

Hydrolytic activity of the catalytically-active, domain-stapled mutant was determined as described in Section 2.3.25. The S243C, S318C mutant showed significantly diminished hydrolytic activity (Figure 5.6) with a specific activity value of 847 ± 68 U/min/μM (enzyme), which is a reduction of 64% when compared to wild-type MMP-1.

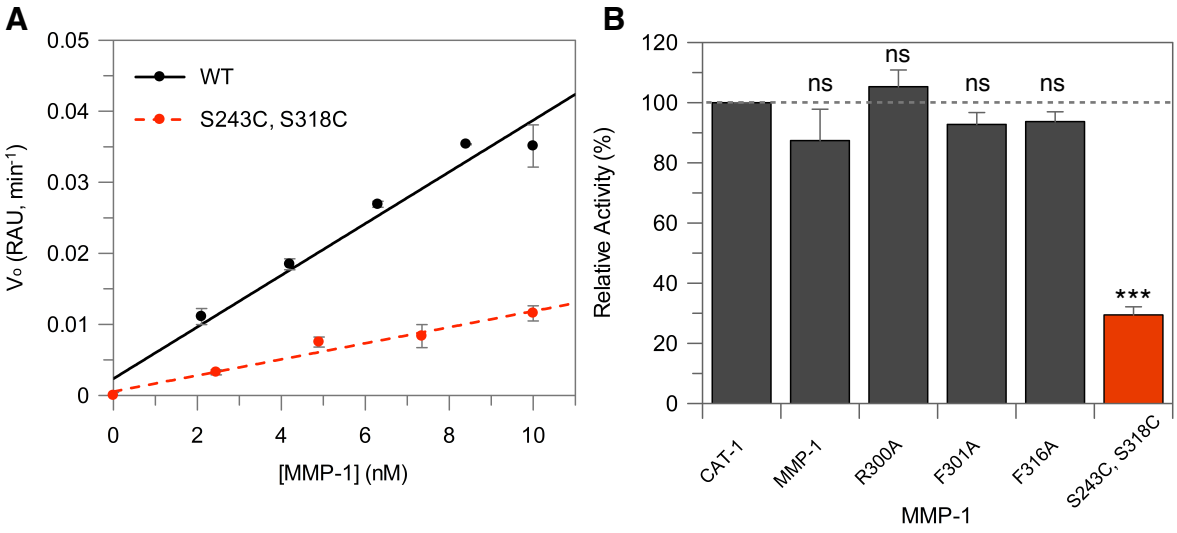



**Figure 5.6: Hydrolytic activity of MMP-1 S243C, S318C.** (**A**) Plot depicting the dose-dependence of reaction rates with MMP-1 S243C, S318C using an MMP-1 catalysed thiopeptolide substrate. Data are mean values of three independent experiments and error bars indicate the SD. (**B**) Plot of relative activity (%) compared to CAT-1 (*horizontal dashed line*). *P*-values were determined by comparing hydrolytic activity of mutant proteins with CAT-1. Statistical analysis was performed using an unpaired two-tailed Student *t*-test (df = 4) and statistically significant *P*-values ($p<0.05$) are shown. The asterisk indicates statistical significance with a *P*-value of $p<0.001$ (***). *P*-values greater than or equal to 0.05 are not considered statistically significant (*ns*).

### 5.3.7 Preliminary crystallisation of MMP-1* S243C, S318C

Crystallisation trials were set up in an attempt to produce an atomic resolution model of the domain-stapled mutant. Small crystals formed after 3 months in a buffer containing 0.1 M succinic acid, pH 7.0 with 15% (w/v) PEG 3350 as a precipitant (Figure 5.7). Current attempts to optimise these conditions have, as yet, been unsuccessful.

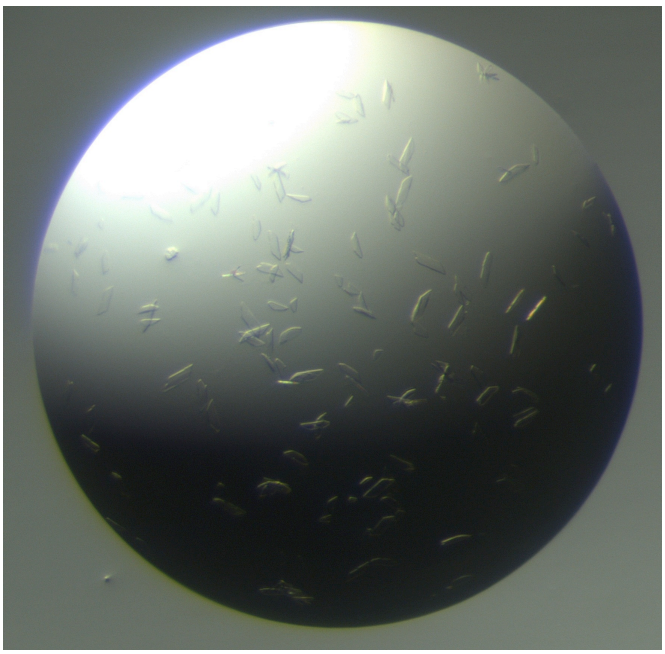


**Figure 5.7: Preliminary crystallographic trial of MMP-1* S243C, S318C.** Conditions: 0.1 M succinic acid, pH 7.0, 15% (w/v) PEG 3350 (200 nl protein at 5 mg/ml with 2 µl precipitant).

## 5.4 Conclusions

Site-directed mutagenesis was successfully used to introduce two serine to cysteine substitution mutations at position 243 in the CAT domain and position 318 in the HPX domain of the pET-3a: proMMP-1* construct. This construct was then transformed into *E.coli* BL21-CodonPlus (DE3)-RIPL cells. Despite comparable WT expression levels after treatment with IPTG, the double mutant appeared to be susceptible to increased aggregation when using the standard proMMP-1* refolding protocol. However, sufficient purified material was obtained and appeared free from soluble disulphide-bonded multimers. Furthermore, global mature enzyme structure appeared unaffected by the double mutation, with no notable increases in hydrodynamic radii as shown by analytical gel filtration.

Reduction in the $R_g$ and $D_{max}$, derived from Guinier and P(*r*) analysis, indicated that in solution MMP-1* S243C, S318C possessed a structure more compact than that of the equivalent wild-type protein. Furthermore, *ab initio* modelling produced a molecular envelope that appeared to lack the regions of density, which are congruent with a dynamic extended state (See Chapter 4). Instead, the data is more consistent with the structural parameters expected for the crystal structure so the mutations have indeed stapled the domains together and prevented domain dislocation.

Limited proteolysis with CAT-3 and CAT-1 was used to determine zymogen activation potential and linker susceptibility, respectively. Equal proteolytic susceptibility was seen when CAT-3 was applied, resulting in enzyme maturation and almost identical digestion profiles for both zymogen mutants, F301A and S243C, S318C. However, the S243C, S318C mutant exhibited significant resistance to hydrolysis of the CAT-HPX interdomain linker when treated with CAT-1. Shielding of the linker region from proteolysis suggests existence of a compact global structure conferred by formation of the interdomain disulphide bond. In addition, it also suggests that, in the WT enzyme, interdomain autolysis only occurs when the enzyme is in its dislocated state. Therefore, the structural basis for linker hydrolytic resistance was investigated using SAXS.

Characterisation of THP-binding by the domain-stapled MMP enzyme was completed using an SPR-based assay. Interestingly, binding by the zymogen was abolished. Enzyme maturation restored some binding activity, but a significant decrease in affinity was still observed (~10-fold). Importantly, neither of the mutated residues appear to be directly involved in substrate recognition since neither residue was detected during a HD/MS study using MMP-1* incubated with a THP substrate (Lauer-Fields et

al., 2009), and alanine substitution of serine 318 had only subtle effects on THP-binding by HPX-1 (See Chapter 3). Therefore, the deleterious effects observed here would again suggest that domain mobility is necessary for exosite exposure and collagen recognition.

Whilst the presence of the HPX domain is an absolute requirement for MMP collagenolysis, cleavage of simple peptides requires only the CAT domain. This was evident when testing the domain-dislocated mutants, all of which showed unimpaired hydrolytic activity compared to the WT enzyme and the CAT domain alone (See Chapter 4). Surprisingly, the mature, domain-stapled revertant displayed compromised hydrolytic activity (~30%) against a peptolide substrate. Previously, inhibitor-induced conformational changes in the catalytic domains of MMP-2, -3, -8 and -13 have been observed (Chen et al., 1999; Feng et al., 2002; Zhang et al., 2000) thereby suggesting that active-site plasticity plays a key role in recognition and hydrolysis of diverse substrates. Therefore, it is possible that introduction of the interdomain disulfide bond has conferred increased rigidity to the CAT domain impacting on the S1' pocket and/or the flexibility of the S1' loop. This observation may have implications for future inhibitor design, whereby MMP hydrolytic activity could be attenuated by targeting long-range interactions that impact active-site plasticity.

# CHAPTER 6: ROLE OF PRO-HPX INTERFACE CONTACTS IN DOMAIN MOBILITY AND COLLAGEN RECOGNITION

## 6.1 Introduction

The crystal structure of human latent and mature MMP-1 reveal substantial global conformational rearrangements occur upon removal of the inactivating PRO-1 domain (Figure 6.1) during enzyme maturation (Jozic, Bourenkov et al. 2005, Iyer, Visse et al. 2006).

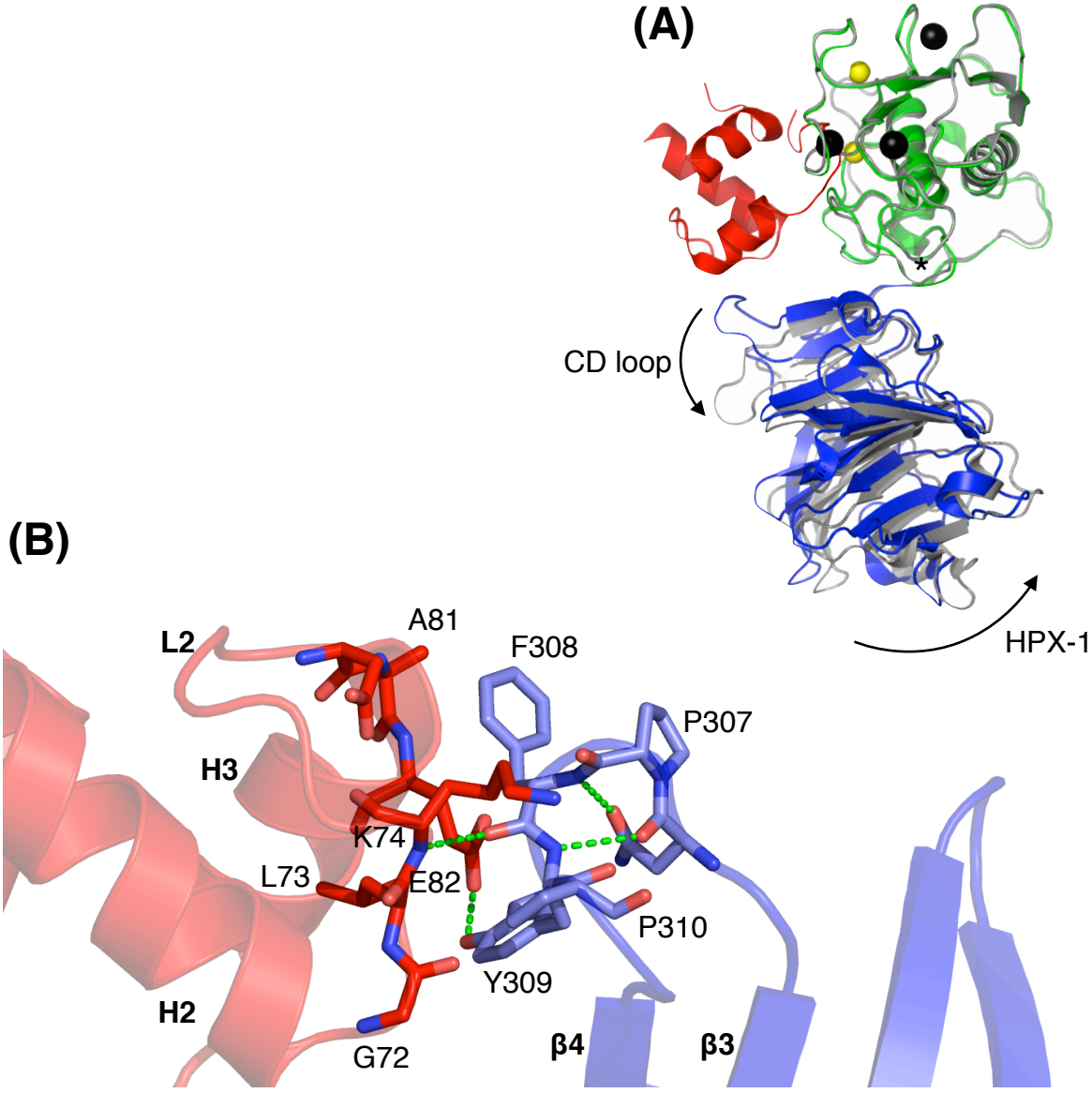



**Figure 6.1: Interdomain interactions of proMMP-1.** PRO-1, CAT-1 and HPX-1 domains are shown in *red*, *green* and *blue* respectively. (**A**) HPX-1 domain rearrangement upon activation. Molecules have been superimposed using their catalytic domains (backbone atomic RMSD 0.815 Å). The curved arrow indicates the relative movement between pro- and active MMP-1. An *asterisk* marks the position of residues RWTNNFREY (**B**) Residues involved in the PRO-HPX interaction are shown as *sticks*. Hydrogen bonds are shown as *green dotted lines*. and the residues are shown as *stick models* in atom type. Secondary structure elements are indicated by *bold* type. Coordinates taken from PDB depositions (1SU3 and 2CLT).

The zymogen structure is restrained by predominantly hydrophobic PRO-HPX contacts that may keep the pro-enzyme in a 'closed' configuration, a state considered unable to bind its collagen substrate (Iyer, Visse et al. 2006). These interdomain contacts occur between PRO-1 residues in loop 2 and helix 3, and residues in the loop between β sheets C and D on blade 1 of the HPX domain. Propeptide removal disrupts these interdomain contacts, unmasking residues Phe308, Tyr309 and Pro310, which undergo large shifts in Cα position (by 16, 12 and 9 Å, respectively) when compared with mature MMP-1 (Iyer, Visse et al. 2006). Correspondingly, a shift is also observed for porcine MMP-1 (Figure 6.2), although much larger at 38, 27 and 24 Å, respectively (Li, Brick et al. 1995, Jozic, Bourenkov et al. 2005).

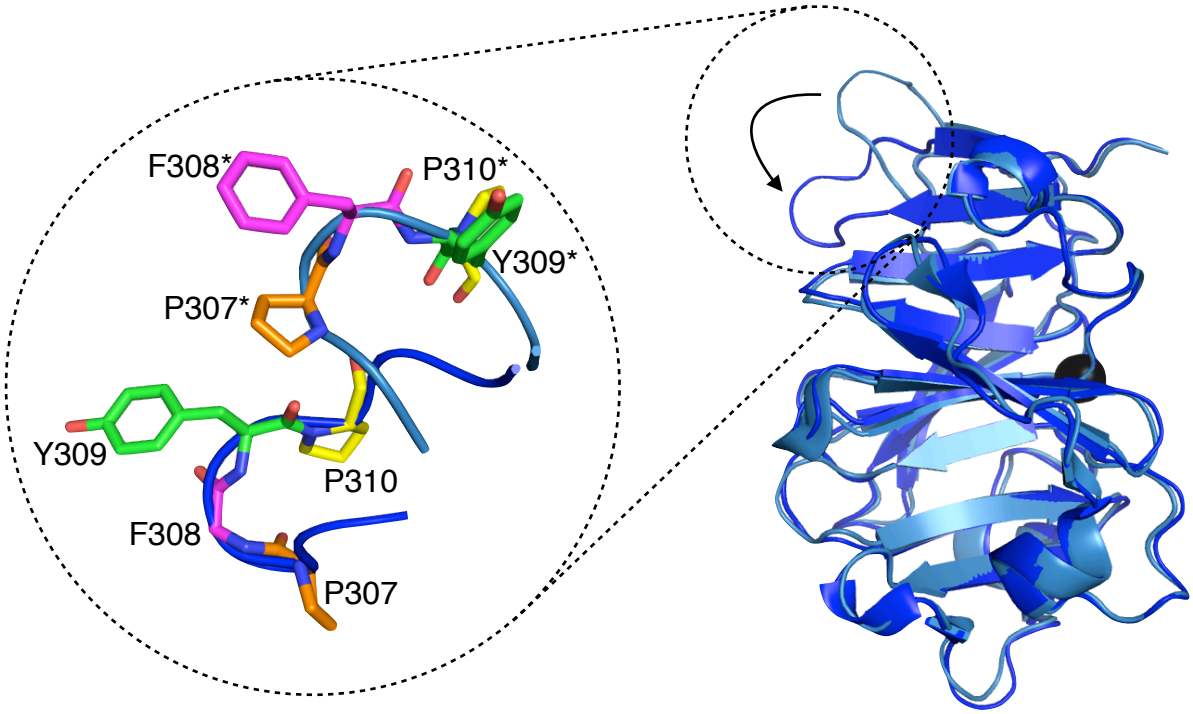



**Figure 6.2: Conformational shift of the blade 1 CD loop region of HPX-1 upon activation.** The HPX-1 of proMMP-1 (PDB accession code 1SU3) and mature MMP-1 (PDB accession code 2CLT) is shown in *light blue* and *dark blue,* respectively. Inset: residues 307-310 are displayed as *sticks*; *orange* P307; *magenta* F308; *green* Y309 and *yellow* P310. Zymogen residues are denoted by an asterisk (*).

The physiological relevance of the PRO-HPX domain interactions is currently unknown. One might speculate that the location of the PRO domain in the catalytic cleft is itself sufficient to prevent hydrolysis. However, these interactions may confer other properties. These could include stabilisation of the PRO domain thus blocking substrate binding and preventing low-level collagen proteolysis, or perhaps restrained domain mobility reducing linker susceptibility to proteolysis and/or exposure

of the collagen binding site, thought to include CAT-1 residues RWTNNFREY (Chung, Shimokawa et al. 2000), which are necessary for collagenolysis (Figure 6.1A).

The role of these PRO-HPX interdomain contacts will be investigated using site-directed mutagenesis of residues in the CD loop to disrupt these interactions, in both the full-length enzyme and HPX-1 protein. The global shape and relative domain orientation of recombinant zymogen proteins will be characterised by SAXS. Substrate binding activity of proMMP-1*, MMP-1* and HPX-1 mutants will be quantified with SPR.

## 6.2 Materials and Methods

The methods necessary to investigate the role of PRO-HPX interdomain contacts have previously been described in Chapters 2-4. Oligonucleotide design and preparation was undertaken as described in Section 2.3.4. Site-directed mutagenesis of target residues was completed as described in Section 2.3.5. Substrate-binding was assayed using SPR as described in Section 3.2.2.2. Low resolution solution structural analysis of recombinant proteins was completed using SAXS as described in Section 4.2.2 and 4.2.3.

## 6.3 Results

### 6.3.1 Mutagenesis

Site-directed mutagenesis of the pET-3a proMMP-1* and HPX-1 plasmid constructs was successfully completed with mutagenic efficiency levels consistent with previous observations (Section 4.3.1).

### 6.3.2 Protein Production

Mutant constructs were transformed into *E. coli* expression strain BL21-CodonPlus (DE3)-RIPL and screened for IPTG-induced recombinant protein expression. Proteins of ~52 kDa and ~24 kDa were expressed, which is consistent with the expected mass predicted for proMMP-1* and HPX-1, respectively.

All target proteins were found to deposit into bacterial inclusion bodies, and were therefore solubilised, refolded and purified in accordance with standard procedures (See Chapter 2). During the refolding procedure, the HPX-1 F308A mutant and all P310A protein forms precipitated leaving insufficient material for further analysis. For all other proteins produced using the standard expression and purification procedure, sufficient milligram quantities were obtained (Table 6.1).

Recombinant proMMP-1* proteins were treated with APMA and CAT-3 as described in Section 2.3.21, and all showed complete enzyme maturation within a four hour incubation period concomitant with an appropriate reduction in molecular weight of ~10 kDa as shown by SDS-PAGE analysis (data not shown). Activated mutant MMP-1* proteins were purified by gel filtration chromatography and eluted at a position consistent with that of WT MMP-1*.

Table 6.1: Purification of recombinant HPX-1 and proMMP* proteins[ab]

| Protein | Fraction | Total Volume (mL) | Protein Concentration (mg/mL) | Total Protein (mg) | Recovered Protein (%) |
|---|---|---|---|---|---|
| HPX-1 P307A | Solubilised Inclusion Bodies | 20 | 5.43 | 108.6 | 100 |
| | Post-CEX | 30 | 0.42 | 12.6 | 11.6 |
| HPX-1 Y309A | Solubilised Inclusion Bodies | 20 | 5.39 | 107.8 | 100 |
| | Post-CEX | 30 | 0.33 | 9.9 | 9.18 |
| proMMP-1* P307A | Solubilised Inclusion Bodies | 20 | 11.82 | 236.4 | 100 |
| | Post-CEX | 30 | 0.78 | 23.4 | 9.89 |
| | Post-GF | 25 | 0.58 | 14.51 | 6.14 |
| proMMP-1* F308A | Solubilised Inclusion Bodies | 20 | 6.393 | 127.86 | 100 |
| | Post-CEX | 30 | 0.582 | 17.46 | 13.66 |
| | Post-GF | 25 | 0.25 | 6.26 | 4.89 |
| proMMP-1* Y309A | Solubilised Inclusion Bodies | 20 | 7.87 | 157.4 | 100 |
| | Post-CEX | 30 | 0.44 | 13.2 | 8.39 |
| | Post-GF | 25 | 0.32 | 8.0 | 5.08 |

[a] from 500 mL of culture

b protein concentration was determined spectrophotometrically ($A_{280}$) using the molar extinction coefficient values obtained from ProtParam analysis (Gasteiger, Gattiker et al. 2003).

### 6.3.3 Functional analysis by SPR

Equilibrium analysis was used to characterise the binding affinity of HPX-1 mutants (P307A and Y309A), proMMP-1* mutants (P307A, F308A and Y309A) and MMP-1* mutants (P307A, F308A and Y309A) for immobilised α1(I)772-787 THP. Triplicate injections were performed using a suitable range of analyte concentrations (0.2 μM-85 μM) and affinities were determined as described in Section 3.2.2.3.

### 6.3.3.1 Binding of HPX-1 proteins to immobilised THP

HPX-1 mutants P307A and Y309A bound to α1(I)772-787 THP with apparent $K_d$ values of 27.4 ± 0.7 μM and 18.1 ± 0.3 μM, respectively (Figure 6.3). Alanine substitution of both these residues clearly diminished the THP substrate binding potential of HPX-1. Interestingly, P307A bound with a 3-fold reduction in $K_d$ value, which approximates that of F301A (Chapter 3). The observed reduction in binding implies that both CD-loop mutants Y309A and P307A may play a role in collagen recognition.

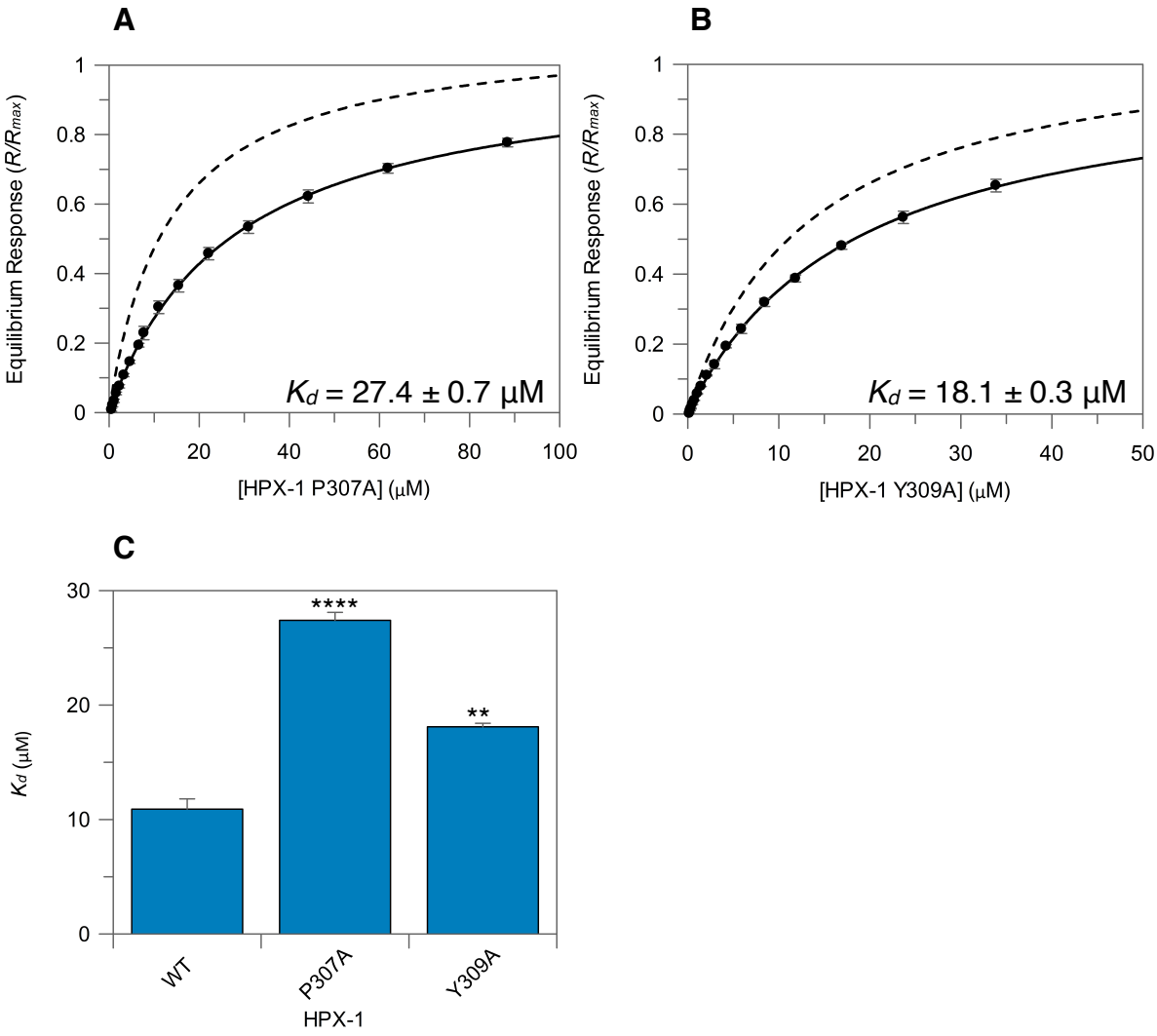



**Figure 6.3: Equilibrium binding of HPX-1 mutants to α1(I)772-787 THP.** Hyperbolic binding curves from equilibrium SPR data (**A**) and (**B**). Binding of wild-type HPX-1 is shown for comparison (*dashed line*). (**C**) Plot of $K_d$ values for binding to immobilised THP. Error bars indicate the S.D. from triplicate measurements. *P*-values were determined by comparing binding affinities of mutant proteins with WT HPX-1. Statistical analysis was performed using an unpaired two-tailed Student's *t*-test (df = 4) and statistically significant *P*-values ($p<0.05$) are shown. The asterisk indicates statistical significance with *P*-values of $p<0.05$ (*), $p<0.01$ (**), $p<0.001$ (***) and $p<0.0001$ (****). *P*-values greater than or equal to 0.05 are not considered statistically significant (*ns*).

### 6.3.3.2 Binding of proMMP-1* proteins to immobilised THP

In contrast, mutations within the HPX CD-loop of proMMP-1* had moderate effects on THP binding when compared with the WT enzyme. The P307A, F308A and Y309A zymogens bound α1(I)772-787 THP with apparent $K_d$ values of 8.1 ± 0.6 μM, 10.4 ± 2.4 μM and 12.9 ± 0.2 μM, respectively. (Figure 6.4). Alanine substitution appeared to impact THP binding affinity for both F308A and Y309A but the greatest effect was observed for the Y309A mutation, which reduced binding by 75%. However, the reduction in binding observed for HPX-1 P307A was masked in the equivalent proMMP-1* mutant, which most closely exhibited proMMP-1* THP binding activity.

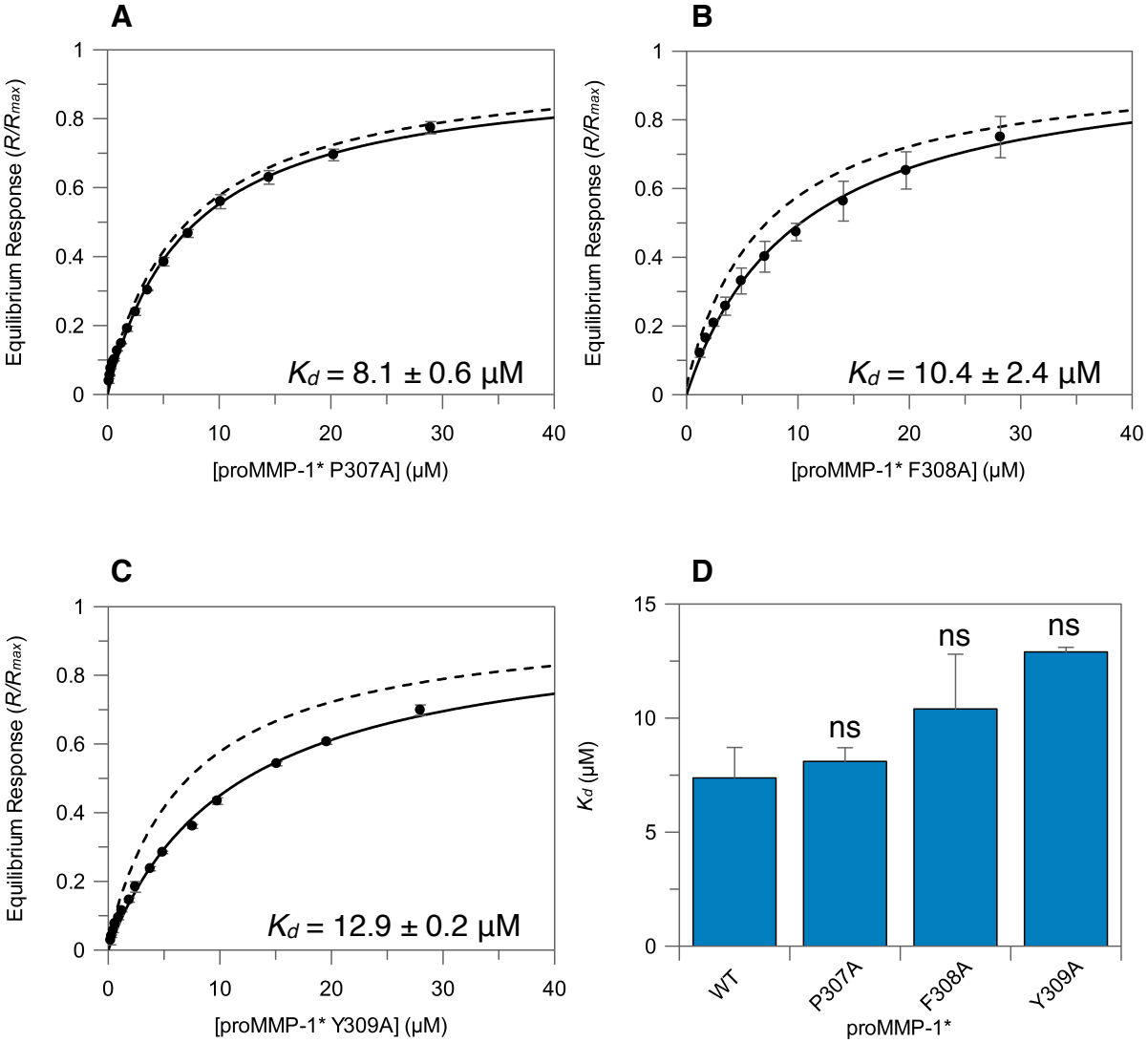



**Figure 6.4: Equilibrium binding of proMMP-1* mutants to α1(I)772-787 THP.** Hyperbolic binding curves from equilibrium SPR data (**A**), (**B**) and (**C**). Binding of proMMP-1* is shown for comparison (*dashed line*). (**D**) Plot of $K_d$ values for binding to immobilised THP. Error bars indicate the S.D. from triplicate measurements. *P*-values were determined by comparing binding affinities of mutant proteins with proMMP-1*. Statistical analysis was performed using an unpaired two-tailed Student's *t*-test (df = 4) to define statistically significant *P*-values ($p<0.05$). *P*-values greater than or equal to 0.05 were not considered statistically significant (*ns*).

### 6.3.3.3 Binding of MMP-1* proteins to immobilised THP

As anticipated, PRO domain removal appeared to increase affinity for the immobilised THP when compared to proMMP-1* mutant forms. The P307A, F308A and Y309A MMP-1* proteins bound α1(I)772-787 THP with apparent $K_d$ values of 0.9 ± 0.4 μM, 2.3 ± 0.05 μM and 1.2 ± 0.02 μM, respectively (Figure 6.5). Again, THP binding by the P307A mutant was not significantly impacted by alanine substitution, suggesting that the presence of the CAT domain is sufficient to restore function for this mutant. Binding activity for Y309A was reduced by 34% , while F308A exhibited the largest reduction in $K_d$ (~1.5 fold).

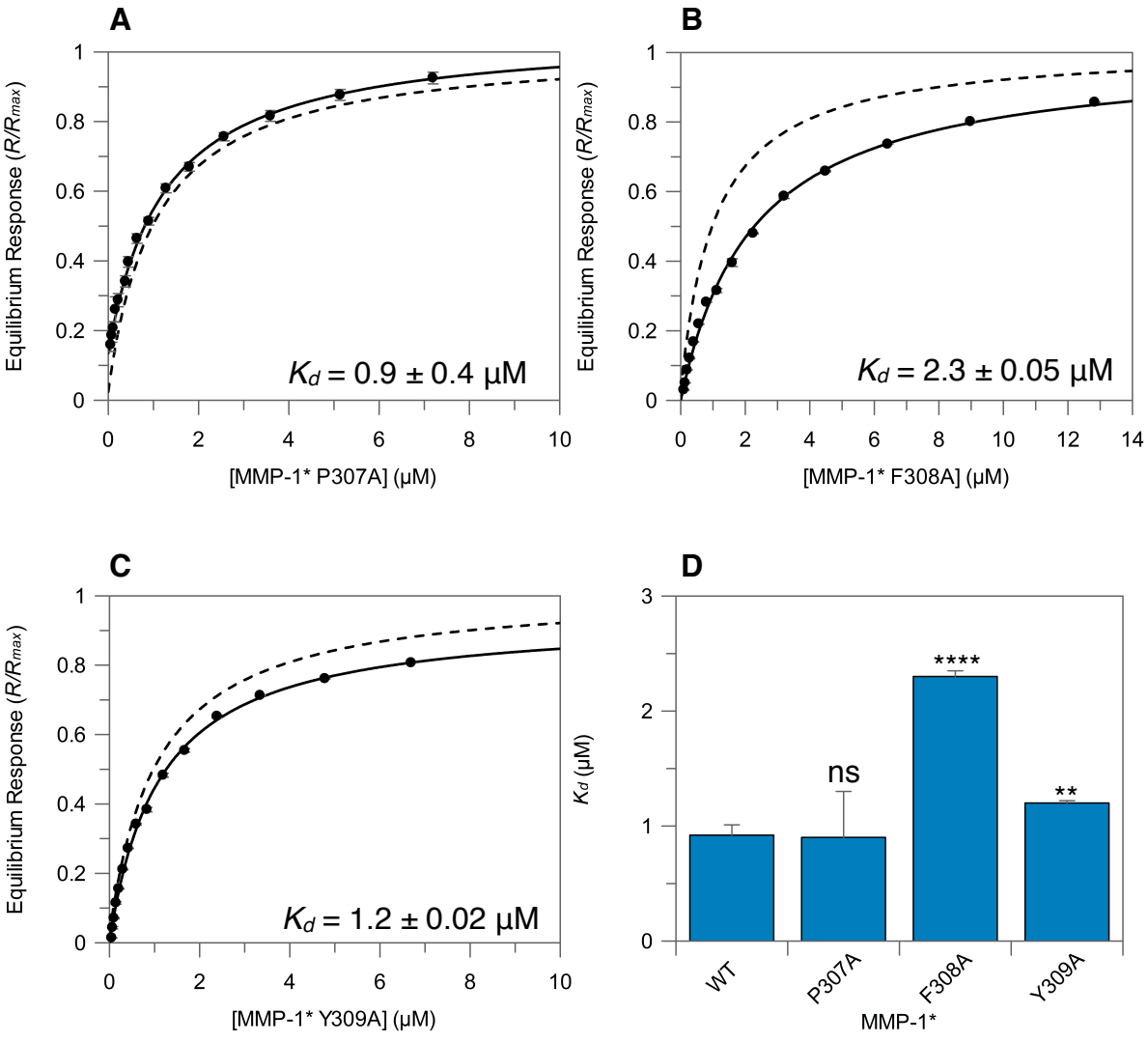



**Figure 6.5: Equilibrium binding of MMP-1* mutants to α1(I)772-787 THP.** Hyperbolic binding curves from equilibrium SPR data (**A**), (**B**) and (**C**). Binding of MMP-1* is shown for comparison (*dashed line*). (**D**) Plot of $K_d$ values for binding to immobilised THP. Error bars indicate the S.D. from triplicate measurements. *P*-values were determined by comparing binding affinities of mutant proteins with MMP-1*. Statistical analysis was performed using an unpaired two-tailed Student's *t*-test (df = 4) and statistically significant *P*-values ($p<0.05$) are shown. The asterisk indicates statistical significance with *P*-values of $p<0.05$ (*), $p<0.01$ (**), $p<0.001$ (***) and $p<0.0001$ (****). *P*-values greater than or equal to 0.05 are not considered statistically significant (*ns*).

### 6.3.4 Structural Analysis of proMMP-1* CD-loop mutants

SAXS was used to determine the structural impact of mutating on residues involved in the PRO-HPX interdomain interaction. SAXS experiments were carried out as described in Section 4.1.1.1 and 4.1.1.2.

The $R_g$, from Guinier analysis, was calculated at 35.8 ± 0.7 Å, 33.2 ± 1.0 Å and 32.0 ± 1.7 Å for proMMP-1* P307A, F308A and Y309A, respectively (Figure 6.6). Only P307A showed significant divergence from proMMP-1* with an increase in $R_g$ by 8%. The asymmetric pair distribution, P(*r*), plots shown in Figure 6.6 indicate that all mutants possess a similar structure to that of the hydrolytically-impaired WT enzyme. The notable observation is a reduction of 4-5 Å in the maximal dimension ($D_{max}$) value in all cases when compared to proMMP-1*. However, despite being a user-defined parameter, each $D_{max}$ value was considered an excellent solution using GNOM (Svergun 1992) and remained consistent for the three mutated proteins.

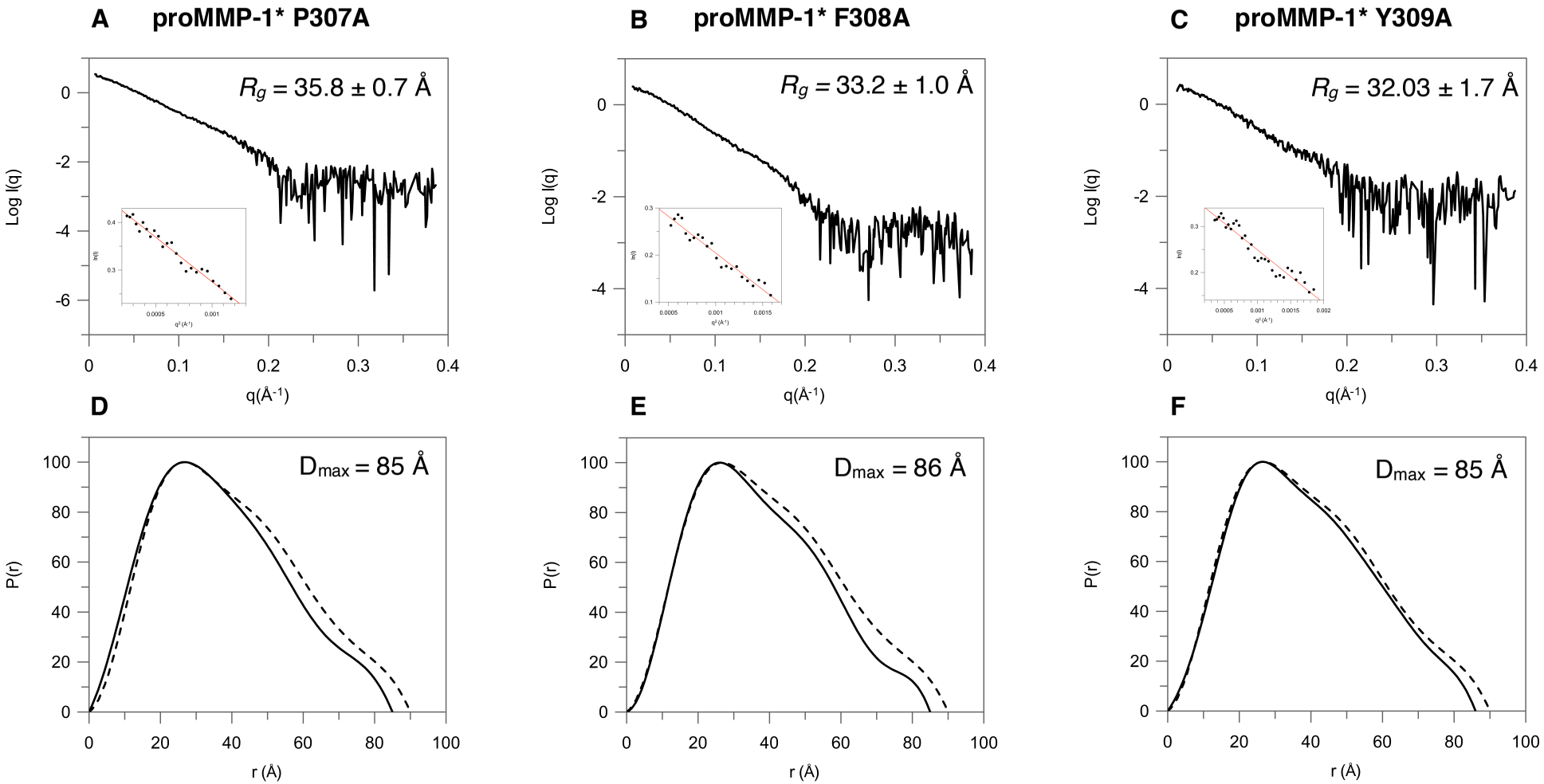



**Figure 6.6: SAXS analysis of proMMP-1* mutants.** (**A**, **B** and **C**) Experimental SAXS scattering profile of mature proMMP-1* mutants after solvent subtraction. Guinier plot of the low-angle region limited by $qR_g$<1.3 (*Inset*). (**D**, **E** and **F**) P(r) function generated by GNOM (Svergun 1992). The proMMP-1* P(r) distribution plot is shown for comparison (*dashed line*).

## 6.4 Conclusions

Mutant pET3a:HPX-1 and proMMP-1* constructs were successfully produced by site-directed mutagenesis and transformed into *E. coli* BL21-CodonPlus (DE3)-RIPL cells. Comparable expression levels with WT proteins were observed in all cases. The standard refolding and purification protocols, devised previously (Chapter 2), resulted in sufficient homologous monodisperse samples of all mutant proteins with the exception of HPX-1 F308A and P310A proteins, which like HPX-1 mutants D299A, F316A and Q352A (Chapter 3), precipitated during refolding.

Proline residues possess unique structural properties as a result of restricted rotation around the N-Cα bond, which precludes adoption of many protein main-chain conformations. Therefore, substitution of proline residues may have deleterious effects on protein folding and/or stability. Consequently, alanine substitution at position 307 may have reduced rigidity in the CD loop between β strands C and D, while disrupting the local hydrogen bond network (Figure 6.1B), and appears to be a critical residue for HPX stability.

The significant reduction in THP-binding observed for HPX-1 P307A may also suggest deleterious structural effects. However, this effect is neutralised in the full-length enzyme. SAXS studies indicate that P307A is similar in shape while exhibiting a larger $R_g$ when compared to proMMP-1*. Yet any structural diversity has virtually no impact on THP-binding in the full-length mutant, which displays equivalent functional WT binding activity. This suggests that P307A has no role in collagen recognition.

In the proMMP-1* crystal structure F308 and Y309 residues are involved in hydrogen-bonding interactions with PRO domain residues K74 and E82, interactions resulting in the 'closed' zymogen configuration (Jozic, Bourenkov et al. 2005). In contrast, the $R_g$ values for F308A and Y309A proMMP-1* mutants obtained using SAXS indicate no significant structural alterations. This seems incongruous with a model involving PRO domain stabilisation via the HPX CD loop. The role of the CD loop appears better attributed to a role in collagen recognition with THP-binding potential reduced in both Y309A and F308A full-length mutants. This observation is supported by a previous study that reported reduced deuterium incorporation in HPX-1 residues 302-316 when incubated with a triple helical peptide substrate (Lauer-Fields, Chalmers et al. 2009). Furthermore, a recent published structure of MMP-1* bound to a collagen peptide showed F308 and Y309 made van der Waals contacts with two α chains. However, in that study a triple CD loop MMP-1 mutant retained 76% of WT

collagenolytic activity  suggesting these residues are not critical for collagen breakdown (Manka, Carafoli et al. 2012).

# CHAPTER 7: DISCUSSION AND FUTURE PERSPECTIVES

## 7.1 Background

The work described herein aimed to explore MMP-1 collagen recognition and the postulated flexible state of MMP-1, with the intention to facilitate understanding of the collagenolytic mechanism. Initially, this required development and optimisation of growth and purification strategies resulting in protein yields sufficient for structural and functional characterisation (Chapter 2).

Using an extensive program of mutagenesis combined with SPR and SAXS, this study has confirmed the presence of a, previously uncharacterised, collagen-binding exosite within the HPX domain (Chapter 3). SAXS experiments have highlighted the discrepancy between MMP-1 solid and solution states and demonstrate, for the first time, solution-state structural alterations upon zymogen activation (Chapter 4). Furthermore, this study has explored the relevance of these interdomain interactions in both proteolytic susceptibility and collagenolytic activity (Chapter 4, 5 and 6).

## 7.2 Potential roles for MMP-1 interdomain flexibility

Originally, comparison of the mature MMP-1 crystal structure with that of proMMP-1 revealed extensive global structural rearrangements, which were posited as a contributory factor in collagen recognition and degradation. In the case of MMP-1, this study confirms that enzyme maturation is accompanied by domain separation. This observation is supported by the recent results from Bertini and co-workers, who demonstrated that mature MMP-1 undergoes transient domain separation concomitant with exposure of domain-domain interface residues using a paramagnetic probe (Bertini, Fragai et al. 2009). Other MMP family members have been shown to adopt an array of conformations as observed in SAXS, AFM, NMR and modelling studies (Rosenblum, Cohen et al. 2007, Bertini, Calderone et al. 2008, Diaz and Suarez 2012) suggesting flexibility is not limited to collagenases, but is a general property of MMPs. However, it is unknown what role these extended interdomain arrangements play in MMP function.

Perhaps, domain mobility is necessary for locating and accessing sites of action. Intramolecular flexibility might facilitate enzyme migration along fibrils and

ECM components. Mobility of MMP-1 and MMP-14 has been investigated during fluorescence correlation spectroscopy experiments, which showed that MMPs appear to move processively along collagen fibrils (Saffarian, Collier et al. 2004, Collier, Legant et al. 2011). Furthermore, in AFM experiments MMP-9 has been shown to undergo conformational contraction upon binding collagen (Rosenblum, Van den Steen et al. 2010) and is able to move processively along the substrate (Collier, Legant et al. 2011). More recently, single-molecule tracking experiments have shown MMP-1 undergoes lateral diffusion along collagen fibrils without noticeable dissociation at a velocity of $11.8 \pm 0.6$ μm $s^{-1}$ (Sarkar, Marmer et al. 2012). However, the molecular mechanism of such mobility remains to be elucidated. It is also unknown whether proMMP-1 possesses the ability to traverse the triple helical substrate, particularly as the zymogen form has reduced affinity for collagen (~ 10-fold). The proMMP-1 solution structure determined here largely agrees with that of the compact crystal structure. Should this compact state restrict mobility it would likely effect the spatial load distribution of enzyme along the fibril and consequently may add another level of regulation.

Arguably, the primary focus of MMP-1 studies remains focused on interstitial collagen breakdown. It is equally possible that extended conformations play a role in other molecular interactions. There is evidence that the HPX domain is required for maximal TIMP inhibition. In TIMP-1 and -4 inhibition studies, full-length MMP-1 displayed higher association rate ($k_{on}$) and lower $K_i$ values when compared with CAT-1 (Murphy, Allan et al. 1992, Troeberg, Tanaka et al. 2002). The same trend was observed for MMP-3. While structural rearrangements have not been directly shown to impact TIMP binding it is difficult to reconcile an interaction with two multi-domain proteins not involving domain mobility, especially when the CAT-3/TIMP-1 crystal complex (PDB accession code 1UEA) shows the C-terminal end of the CAT-3 polypeptide chain oriented well away from full-length TIMP-1. Equally, adoption of a range of variable structural conformations would likely allow the ability to bind and/or process a diverse substrate repertoire. PAR-1, CD44, integrins, MCPs (monocyte chemoattractant protein) and TGF-β are just a few of the known MMP substrates (McQuibban, Gong et al. 2002, Boire, Covic et al. 2005, Chetty, Vanamala et al. 2012, Kryczka, Stasiak et al. 2012). Many of these molecules are associated with the cell surface and may be assembled in large multi-component complexes. As such, variable inter-conformations would maintain MMP versatility.

It has been postulated that active-site plasticity is necessary for such diverse substrate recognition and processing (Lovejoy, Welch et al. 1999, Moy, Chanda et al.

1999). An unexpected finding from this study was the compromised hydrolytic activity resulting from introduction of an interdomain disulfide bond (MMP-1 S243C, S318C) to confine the enzyme to a compact state. In contrast, the activity of domain-dislocated mutants (MMP-1 R300A, F301A and F316A) was unimpaired. Therefore, promotion of compact states *in vivo* may confer rigidity to the CAT-1 subsite pockets, thereby spatially and temporally modulating activity against preferred substrates. Furthermore, the compact state appears to confer resistance to proteolysis of the inter-domain linker. This too may have significant consequences, by either promoting retention of the HPX domain for specific localisation within the ECM, regulating functions that are independent of catalytic activity and are mediated solely by the HPX domain (i.e. cell signalling pathways (Mantuano, Inoue et al. 2008)) or reducing clearance of MMPs from the ECM. What is certain, the compact state has biological relevance in collagen processing, as complete disruption of the CAT-HPX interface, confirmed by SAXS experiments, drastically reduces (~ 30-fold) collagen binding activity in mature MMP-1.

## 7.3 Collagen recognition and degradation

During the latter stages of this project, two studies were published that provide considerable insight into collagen catabolism by MMP-1. The work conducted by Bertini and co-workers describes a mechanism of collagenolysis that is derived from NMR experimental data of enzyme-substrate complexes, and docking experiments using docking programme HADDOCK (Bertini, Fragai et al. 2012).

Initially, the triple helical peptide structure was confirmed and assigned using $^{13}C,^{15}N$-enriched α1(I)772–786 THP. Full-length enzyme and isolated domains in complex with the THP were subsequently monitored by $^{1}H,^{15}N$ TROSY-HSQC spectroscopy for changes in signal intensity. This determined a notably strong localised interaction between HPX-1 (residues 291-292 and 311-326) with two chains of the THP (analogous to type I collagen residues 782-785), with further points of peptide interaction with active site residues in CAT-1 (residues 160-199 and 216-224). In part, this agrees with some of the data described in this thesis (Chapter 3 and 4). SPR THP binding data confirms reduced affinity for alanine mutated residue V319 (26.4 ± 1.5 μM) in the isolated HPX-1 domain. Furthermore, mutation of F316 in full-length MMP-1 reduces THP affinity 2-fold. However, despite their selection as restraints in the iterative HADDOCK docking process, residues R300 and F301 failed to be identified in the NMR study. Instead Bertini and co-workers used these residues as pseudo-restraints,

along with 233, 241, 243, 247, 250, 271, 272, 300, 301, 316, 318 and 326, which constitute CAT-HPX interface residues and part of the linker, to induce structural reorientation during the docking procedure. The SAXS data presented herein for MMP-1 R300A, F301A and F316A, which highlights elongated solution structural states, would support this rationale. In addition, the linker proteolysis study (Section 5.3.4) supports a model of an enzyme with domains that transiently separate (Figure 7.1A) leading to linker exposure. The collagenolytic mechanism proposed by Bertini et al. (2012) describes a process whereby an extended MMP-1 molecule forms an 'encounter complex' with the THP (Figure 7.1B) as defined by NMR data.

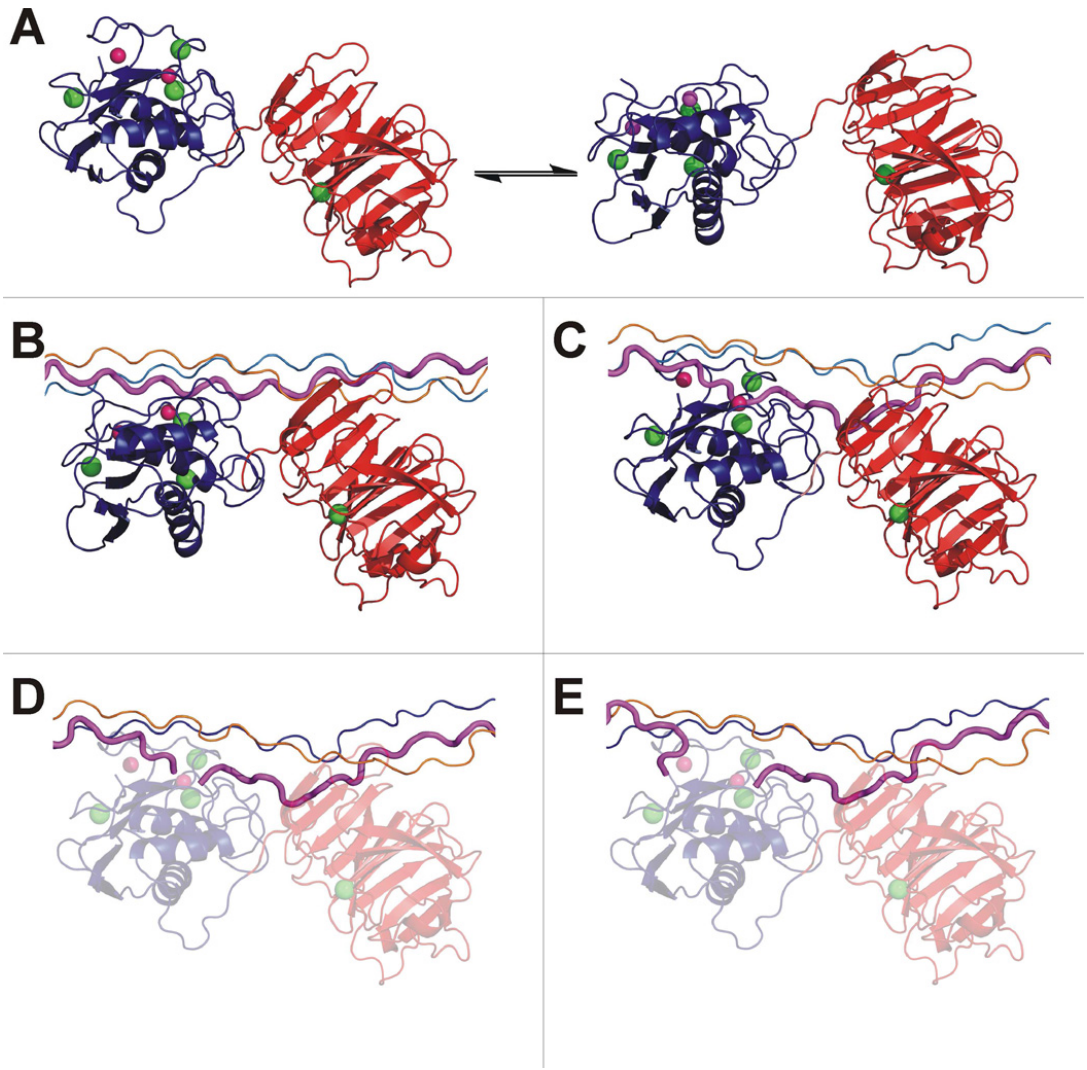



**Figure 7.1: The Bertini Model of collagenolysis.** (**A**) Compact (*left*) and extended (*right*) forms of mature MMP-1 in equilibrium. (**B**) Initially extended MMP-1 binds the triple helix. The HPX domain binds via four amino acids (analogous to type I collagen residues 782-785) and the triple helix is presented to the catalytic domain. (**C**) In the compact conformation, MMP-1 releases one chain (*magenta*), which is positioned in the CAT domain active site for hydrolysis. (**D**) After hydrolysis, both peptide fragments (*C*- and *N*-terminal) are initially bound to the active site. (**E**) The *C*-terminal region of the *N*-terminal peptide fragment is released. (Reproduced from Bertini, Fragai et al. 2012).

Nuclear overhauser effects (NOEs), which are a useful probe of spatial proximity, indicate distortion of the THP helical state in the presence of MMP-1. Therefore, the authors speculate that domain back-rotation to form the compact crystallographic MMP-1 state is required to induce substrate distortion and unwinding in preparation for hydrolysis (Figure 7.1C). It must be stressed that there is no experimental evidence that domain mobility is retained upon enzyme-substrate complex formation, although our data show that permanently separated domain states have negative effects on THP-binding.

The proposed final substrate cleavage steps (Figure 7.1D and E) are the most tenuous, using previously published MMP-12 data (PDB accession code: 2OXZ) as a model for hydrolysis (Bertini, Calderone et al. 2006). This study involved crystallisation of the catalytic domain of MMP-12 in the presence of a simple peptide in order to study the steps of the MMP reaction mechanism. The relevance of this data is questionable; MMP-12 is not a collagenase and therefore is unable to process complex triple helical collagen-like peptides. Furthermore, the study focused on CAT domain interactions only. As such, there is no evidence that involvement of the HPX domain remains in the subsequent chain processing stages of collagenolysis.

The second study performed by Manka and co-workers culminated in the first crystallographic structure of MMP-1 in complex with a THP (PDB accession code: 4AUO), showing that collagenolysis relies on multiple exosite interactions (Manka, Carafoli et al. 2012).

Analysis of a Collagen Toolkit library of triple-helical peptides encompassing the entire collagen II sequence confirmed the importance of leucine residues at positions P1‘ and P10’, thereby validating the choice of THP used in this study (Section 3.1.4), which contains the necessary binding motif (residues 772-787) for MMP-1 recognition.

Using an ELISA assay, the authors confirm MMP-1* binds native type I collagen with an apparent $K_d$ of 0.4 μM. In contrast, CAT-1* showed no detectable binding and HPX-1 bound weakly, thereby illustrating the necessity of cooperative domain interactions for effective collagen recognition. This substantiates the experimental SPR data presented here, which saw an order of magnitude reduction in $K_d$ value for MMP-1* (0.92 ± 0.1 μM) when compared to HPX-1 (10.9 ± 1.0 μM). Moreover, HD/MS experiments highlight regions involved in type I collagen binding (Figure 7.2A) and clearly identify multiple sites that span both domains. Delayed

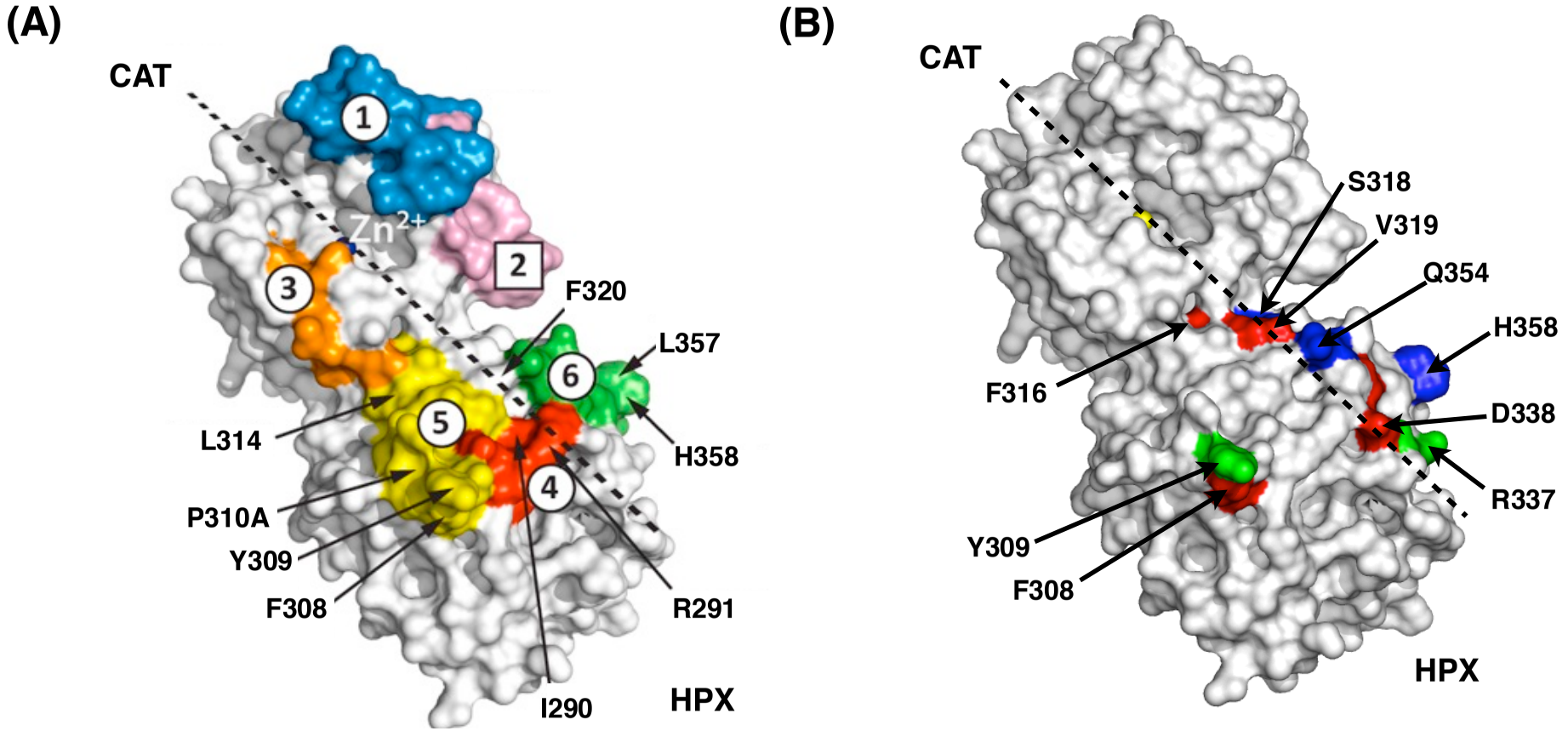



**Figure 7.2: Collagen Binding Site.** (**A**) Type I collagen footprint determined by HD/MS and mutagenesis **(Manka, Carafoli et al. 2012)**, and mapped onto the crystal structure of MMP-1* (PDB accession code: 2CLT). Sites 1 and 3–6 are protected from deuterium incorporation, and site 2 shows enhanced deuterium incorporation upon collagen binding. *Dashed line*, predicted collagen binding direction. (B) Residues of HPX-1 implicated in collagen binding by mutagenesis and assay (Chapters 3 and 6). Mutated HPX-1 residues are colored according to the collagen binding activity of the mutant, *i.e.* V319, and D338 in *red* ($K_d \geq 24$ μM); Y309, R337 and P361 in *green* (16 μM $< K_d <$ 24 μM); and S318, P325, Q354, and H358 in *blue* ($K_d \leq 16$ μM). Zinc ion is depicted as a *yellow sphere*. *Dashed line*, predicted collagen binding direction.

deuterium exchange was observed across the following sites: residues 164-180 (site 1), 224-234 (site 3), 285-296 (site 4), 302-316 (site 5), 349-365 (site 6)). Enhanced H/D exchange was observed for residues 198-212 (site 2). Sites 4 and 5 have been previously identified in similar experiments (Lauer-Fields, Chalmers et al. 2009) and constitute surface regions of HPX-1 blade 1. The residues selected for investigation in this thesis are primarily located in blade 1 but focus on individual contributions to collagen recognition. Reassuringly, most of these residues are found in sites 5 and 6 (Figure 7.2B). However, of the 'ball' residues selected in this study, only F316 (site 5) displayed deuterium exchange. This is not surprising as most residues comprising the 'ball and socket' would be protected from the solvent in the compact state thereby reducing the potential for detection.

Interestingly, site 2 is composed of the sequence RWTNFFREY, previously identified as a critical component of collagenolytic activity (Chung, Yoshida et al. 2004). Deuterium exchange amplification in this region is suggestive of increased exposure and mobility within in the CAT domain. Therefore, site 2 may be another long-range region affected by domain-stapling as introduction of mutation S243C, S318C reduced both hydrolytic activity and THP binding (~10 fold). It may also be suggestive of a mobility mechanism that requires domain separation, either for initial recognition or in later binding events.

Further mutational analysis of candidate residues identified by docking was performed, and collagenolytic activity was tested. The double mutant L357A, H358A showed a marginal increase in relative collagen hydrolysis. This fits with the SPR data presented in Section 3.2.2.2, which also showed a slight decrease in $K_d$ for single mutant H358A. Double mutant I290A/R291A (Lauer-Fields, Chalmers et al. 2009) exhibited the largest reduction, retaining only 4% of activity. A review of this position (Figure 7.2) shows the close proximity of D338. Further analysis of the crystal structure (PDB accession code: 2CLT) indicates the presence of a salt bridge between this residue and R291. Therefore, disruption of this interaction may account for the significant reduction in both affinity ($K_d$ = 24.5 ± 0.9 μM) and subsequent hydrolysis. Hydrolytic activity of mutant F320Y also showed reduced activity (10%). The crystal structure complex (Figure 7.3) indicates a role in forming the S10' binding pocket of HPX-1. However, there is no attempt to assess the structural impact of this mutation, it is therefore possible that mutation has affected positioning of adjacent residue V319, which when mutated is negatively impacted with an increase in $K_d$ (26.4 ± 1.5 μM).

Perhaps most intriguing is the role of CD loop residues. Triple alanine

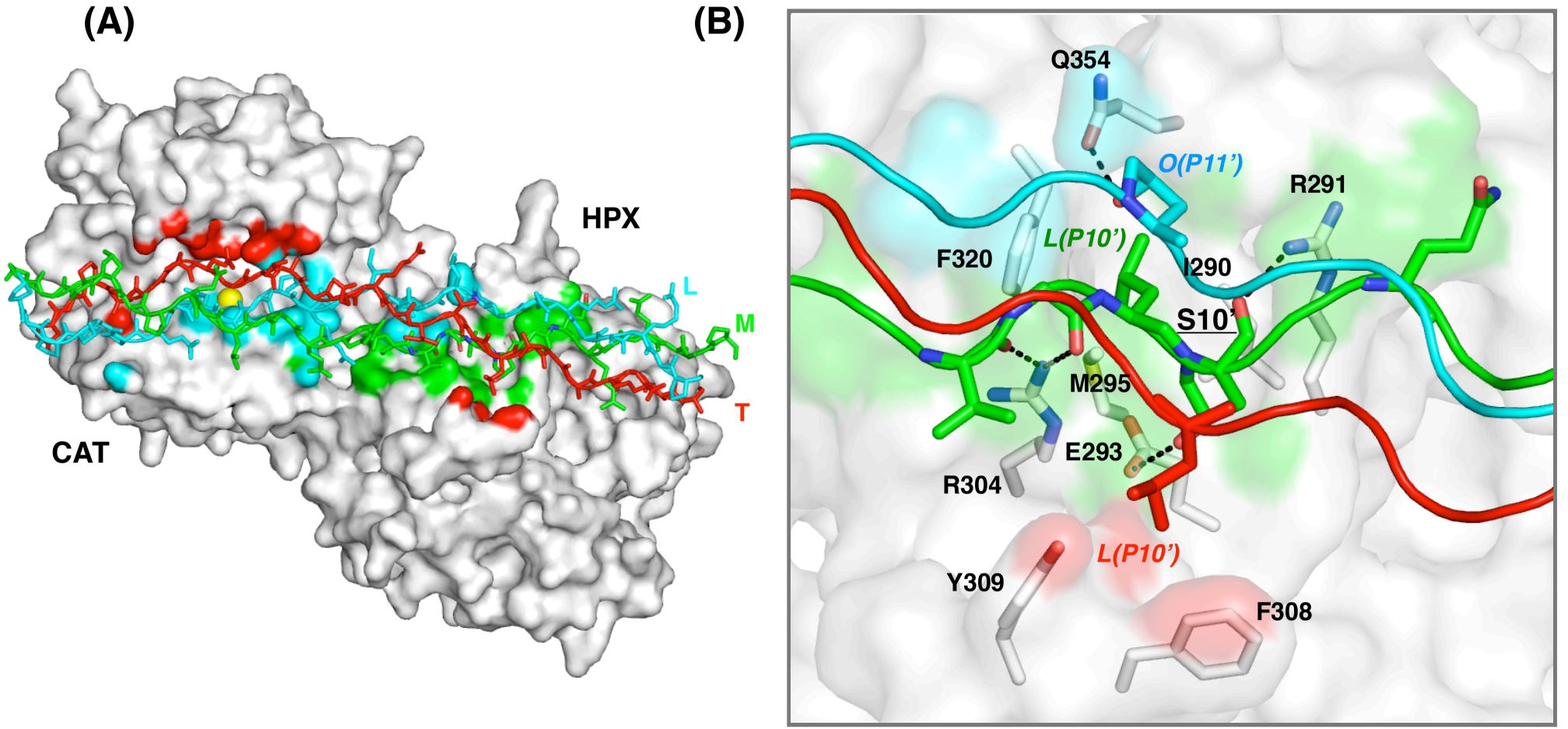



**Figure 7.3: Crystal structure of the MMP-1*–THP complex.** (**A**) MMP-1* is shown as a *gray surface* with the THP chains colored *cyan*, *green* and *red*. Surface areas within 4 Å distance of the leading (**L**), middle(**M**) and trailing (**T**) chains are colored correspondingly. *Yellow sphere*, active-site zinc ion. (**B**) View of the interactions of the THP chains with the HPX-1. *Dashed lines* indicate hydrogen bonds. Residues making enzyme-substrate contact are labeled accordingly. (Adapted from Manka, Carafoli et al. 2012)

substitution of F308,Y309 and P310 only reduces hydrolytic activity by ~25%, yet both F308 and Y309 make hydrophobic contacts with the P10' residue of the trailing strand that is considered a critical binding motif (Figure 7.3B). Furthermore, the largest reduction in $K_d$ for the CD loop residues was observed during SPR experiments of single mutant MMP-1* F308A (~1.5 fold). However, this reduction in affinity was moderate when placed in context of other full-length mutants tested; MMP-1* R300A, F301A and F316A showed reductions in $K_d$ of 3-fold, 30-fold and 3-fold, respectively. Then consider the reduction in type I collagenolysis (13%) displayed by linker residue G272A (Tsukada and Pourmotabbed 2002). This would imply that domain mobility and cooperation, with stabilisation from catalytic domain residues is more important than the hydrophobic interactions provided by the CD loop.

Notably, the wall-forming segment of the CAT-1 S1' pocket (G233-A-L-M-Y-P-S-Y240) is proximal to the 'socket' residues, and a hydrogen bond exists between the G233 main chain carbonyl group and 'ball' residue R300. Potential disruption of this bond may account for the reduced THP affinity observed. Another factor may reside in the impact on Y237, P238, S239 and Y240, which are directly involved in binding the leading chain (Figure 7.4), by MMP-1 domain separation.

The MMP-1*-collagen peptide complex is undoubtedly important, providing for the first time clear structural insights into collagenolysis. However, like the Bertini model described previously, it too poses some controversy. The structure is of a type II collagen homotrimeric peptide bound to MMP-1*. Despite differences in MMP-1 kinetic parameters for different collagen substrates (Table 1.1), presumably the mechanism of binding is similar to type I collagen. The crytallised complex is deemed "non-productive" by the authors as the S1' pocket remains empty with the leucine at P1' 9 Å away. They recognise that a productive mode could be achieved by separation of the CAT-HPX interface, which is consistent with experimental reports of MMP flexibility in solution and the data presented herein. Yet, this was discounted in favour of a proposed mechanism that involves a compact MMP-1* state and axial rotation of the substrate, despite no experimental evidence to support this proposal.

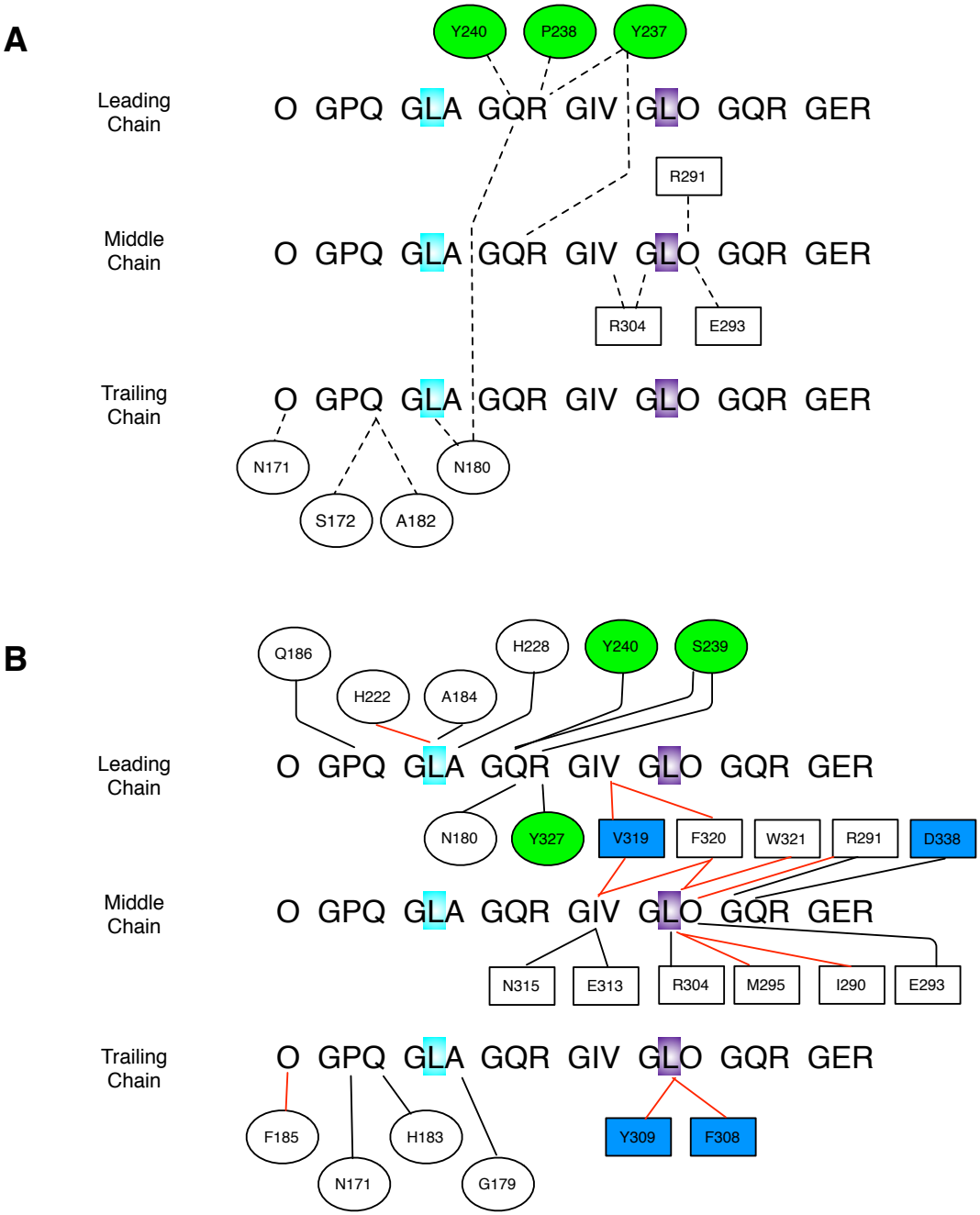



**Figure 7.4: Schematic representation of MMP-1 collagen interactions.** MMP-1 residues are shown as ovals (CAT-1) or rectangles (HPX-1). Collagen subsites P1' and P10' are shown in *cyan* and *purple*, respectively. (A) Hydrogen bonds (*dashed lines*) between MMP-1 and THP. (B) Contacts of 4 Å distance or less (*solid lines*) between MMP-1 and THP. Hydrophobic contacts are emphasized by *red lines*. Residues involved in forming the S1' pocket are shown in *green*. HPX-1 residues with significantly reduced THP affinity upon alanine substitution are shown in *blue* (Adapted from Manka, Carafoli et al. 2012).

The crystallised complex is in a compact structural form, not dissimilar to that shown for MMP-1* S243C, S318C (Chapter 5). Mutagenic modelling of the MMP-1:THP complex reveals a structure also amenable to cysteine substitution at position 243 and 318 (Figure 7.5). However, the data presented here demonstrates, that when trapped in a compact form, MMP-1* S243C, S318C shows a reduction in THP affinity (~10-fold reduction in $K_d$). Arguably, while the compact state plays a role in THP binding due to the negative effects of domain dislocation, the compact form is also detrimental to effective collagen-processing.

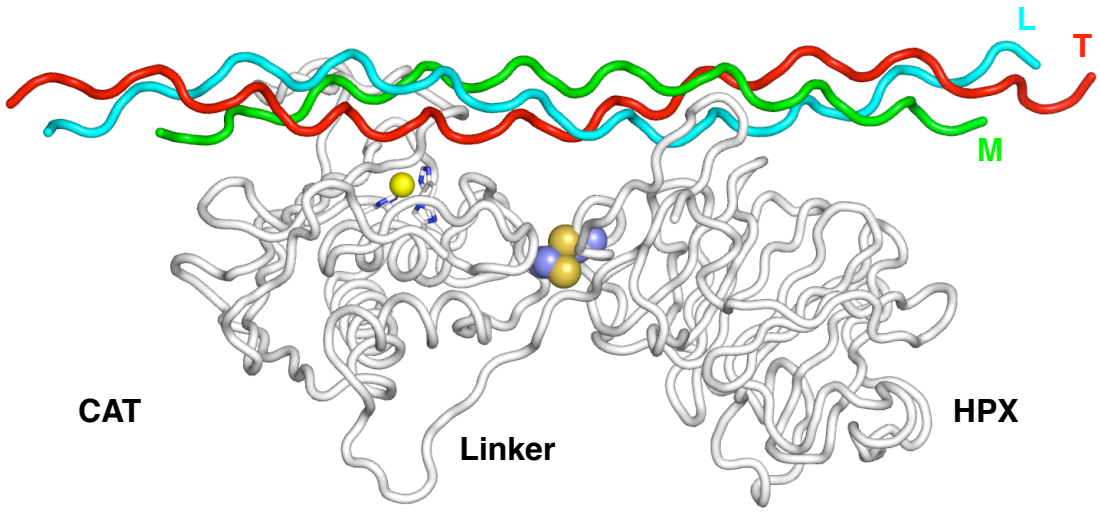



**Figure 7.5: A model of MMP-1* S243C, S318C complexed to a collagen peptide.** The crystallised complex (PDB accession code: 4AUO) is depicted as a cartoon. Note the distance of the trailing chain from the active site zinc atom (*yellow sphere*) with the histidine triad shown as *sticks*. Cysteine substitution at serine positions 243 and 318 has been modelled using the Pymol Molecular Graphics System (version 1.5.0.4.), Schrödinger, LLC. The mutations are shown as *spheres* are coloured according to atom. The upper rim of the catalytic site cleft anchors the middle chain (*green*), and the HPX-1 anchors the leading chain (*cyan*).

It is clear that both studies offer valuable insight into collagen breakdown. Yet, both present conflicting models of collagen recognition and processing. The data presented here would suggest a structurally dynamic enzymatic form exists that requires both an elongated and compact form for effective collagen binding. Therefore, with further study, the mutants used in this thesis could assist elucidation of the correct role of elongated and compact states in collagenolysis.

## 7.4 Proposed future work

Clearly the mechanism of collagenolysis by MMP-1 remains to be fully elucidated. While the structural and functional characterisation of MMP-1, the subject of this thesis, has provided valuable insights into cooperative domain interactions, there remains many outstanding questions that require answers. The following discussion covers many potential avenues of investigation but is by no means an exhaustive list.

Firstly, the structure of the domain stapled mutant requires further validation. The SAXS, proteolysis and SPR data presented here suggest MMP-1* S243C, S318C is a more compact molecule than the wild-type enzyme, with subsequent negative effects on substrate interactions. Ideally, an atomic resolution model of the domain stapled mutant would corroborate the low resolution structural data. At present, preliminary crystallisation has taken place, small crystals have formed but attempts to optimise the crystallisation conditions have, as yet, been unsuccessful. Therefore, an NMR study to study structure and dynamics, like that performed by Bertini et al. (2012), may be more appropriate. The impaired hydrolytic activity may result from CAT domain rigidity. Therefore, differential scanning calorimetry should be used to assess the thermostability of this mutant. Furthermore, cysteine bond reduction using oxidative agents could be used to assess restoration of THP binding and hydrolytic activity.

The work contained herein, looked primarily at the effect of structural flexibility on THP recognition and binding. The actual effect of elongated and compact states on collagen hydrolysis remains to be identified. Utilising mutants that are effectively trapped in either permanently dislocated or compact stapled forms should allow full elucidation of kinetic parameters (on/off rates) and definition of $k_{cat}/K_M$ values for collagen breakdown thereby solving the nature of the initial encounter complex that remains a contentious point in the previously described Bertini and Manka models. This could be complemented with HD/MS experiments in order to explore potential changes in the collagen footprint of these mutants for comparison to the sites identified by Manka and co-workers.

Long-term prospects might include targeting transient domain separation and/or HPX exosites involved in collagen recognition as a potential therapeutic avenue. One of the fundamental issues with MMP inhibitors is the broad spectrum effects as a consequence of high homology in the CAT domain (Reveiwed in Nuti, Tuccinardi et al. 2007). By targeting exosites and/or mobility mechanisms it may be possible to attenuate specific collagenase activity by MMP-1. Small molecule inhibitors have already been

developed against the HPX domain of MMP-14 with significant reductive effects on protumorigenic activity *in vivo* (Remacle, Golubkov et al. 2012). Such an approach would need to include compound generation, selection and screening, determination of inhibiton $K_i$ values and assessments of *in vivo* effects.

Finally, as a variety of studies exist that suggest interdomain flexibility is a general property of MMPs (MMP-2, -9, -12), the methods of experimentation utilised in this study could be applied to other MMP family members like the gelatinases and stromelysins. Alignment of seven MMP sequences (Figure 7.6) show that the arginine residue at position 300 is strictly conserved with a general preference for an aromatic residue at position 301. In MMP-1 R300 interacts with G233 of the wall-forming socket and therefore this may play a role in substrate hydrolysis.

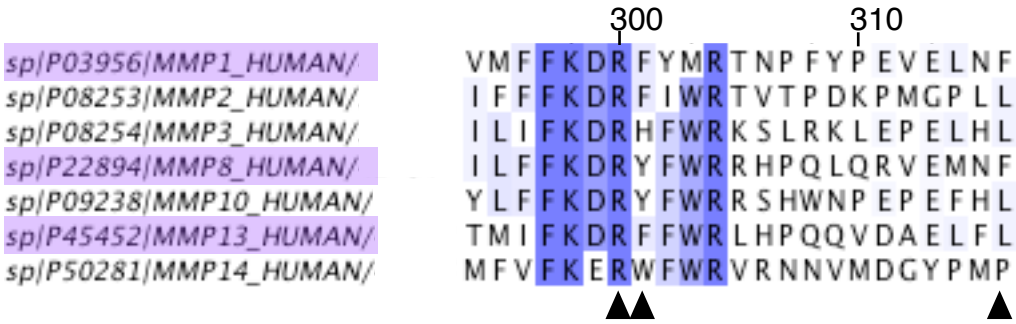



**Figure 7.6: MMP alignment of proposed interface residues.** Residues comprising the MMP-1 'ball' joint are indicated by *black triangles*. Amino acid sequence alignment is coloured with increasing percentage identity (75-100%). Members of the collagenase family are highlighted in *lilac*. Protein sequences were obtained from the Uniprot/Swiss-Prot database (Apweiler, Bairoch et al. 2004), aligned using CLUSTAL Omega (Sievers, Wilm et al. 2011) and rendered using Jalview (Waterhouse, Procter et al. 2009).

In conclusion, the data presented here identifies HPX-1 residues forming a collagen binding exosite, this has since been validated by crystallisation of an MMP-1*-THP complex (Manka, Carafoli et al. 2012). Moreover, the data confirms, the previously postulated, flexible state of MMP-1 and has shown that interfering with this subtle dynamic equilibrium has significant effects on collagen recognition.

# REFERENCES


Agilent Technologies, I. (n.d.). "Quikchange Site-Directed Mutagenesis Kit Instruction Manual." 1-14.

Ahokas, K., J. Lohi, H. Lohi, O. Elomaa, M. L. Karjalainen-Lindsberg, J. Kere and U. Saarialho-Kere (2002). "Matrix metalloproteinase-21, the human orthologue for XMMP, is expressed during fetal development and in cancer." Gene **301**(1-2): 31-41.

Aimes, R. T. and J. P. Quigley (1995). "Matrix metalloproteinase-2 is an interstitial collagenase. Inhibitor-free enzyme catalyzes the cleavage of collagen fibrils and soluble native type I collagen generating the specific 3/4- and 1/4-length fragments." J Biol Chem **270**(11): 5872-5876.

Airola, K., T. Karonen, M. Vaalamo, K. Lehti, J. Lohi, A. L. Kariniemi, J. Keski-Oja and U. K. Saarialho-Kere (1999). "Expression of collagenases-1 and -3 and their inhibitors TIMP-1 and -3 correlates with the level of invasion in malignant melanomas." Br J Cancer **80**(5-6): 733-743.

Alexander, C. M., E. J. Hansell, O. Behrendtsen, M. L. Flannery, N. S. Kishnani, S. P. Hawkes and Z. Werb (1996). "Expression and function of matrix metalloproteinases and their inhibitors at the maternal-embryonic boundary during mouse embryo implantation." Development **122**(6): 1723-1736.

Allan, J. A., A. J. Docherty, P. J. Barker, N. S. Huskisson, J. J. Reynolds and G. Murphy (1995). "Binding of gelatinases A and B to type-I collagen and other matrix components." Biochem J **309 ( Pt 1)**: 299-306.

Allan, J. A., R. M. Hembry, S. Angal, J. J. Reynolds and G. Murphy (1991). "Binding of latent and high Mr active forms of stromelysin to collagen is mediated by the C-terminal domain." J Cell Sci **99 ( Pt 4)**: 789-795.

Andronicos, N. M., E. I. Chen, N. Baik, H. Bai, C. M. Parmer, W. B. Kiosses, M. P. Kamps, J. R. Yates, 3rd, R. J. Parmer and L. A. Miles (2010). "Proteomics-based discovery of a novel, structurally unique, and developmentally regulated plasminogen receptor, Plg-RKT, a major regulator of cell surface plasminogen activation." Blood **115**(7): 1319-1330.

Andronicos, N. M. and M. Ranson (2001). "The topology of plasminogen binding and activation on the surface of human breast cancer cells." Br J Cancer **85**(6): 909-916.

Anthony, L. C., Dombkowski, A. A., & Burgess, R. R. (2002). Using disulfide bond engineering to study conformational changes in the beta'260-309 coiled-coil region of Escherichia coli RNA polymerase during sigma(70) binding. *J Bacteriol, 184*(10), 2634-2641.

Apweiler, R., A. Bairoch, C. H. Wu, W. C. Barker, B. Boeckmann, S. Ferro, E. Gasteiger, H. Huang, R. Lopez, M. Magrane, M. J. Martin, D. A. Natale, C. O'Donovan, N. Redaschi and L. S. Yeh (2004). "UniProt: the Universal Protein knowledgebase." Nucleic Acids Res **32** (Database issue): D115-119.

Arnold, L. H. (2010). Biophysical Characterization of Collagen Binding by the Hemopexin Domain of Matrix Metalloproteinase-1 (MMP-1). (Ph.D.), University of Portsmouth, Portsmouth, UK.

Artym, V. V., Y. Zhang, F. Seillier-Moiseiwitsch, K. M. Yamada and S. C. Mueller (2006). "Dynamic interactions of cortactin and membrane type 1 matrix metalloproteinase at invadopodia: defining the stages of invadopodia formation and function." Cancer Res **66**(6): 3034-3043.

Atkinson, S. J., T. Crabbe, S. Cowell, R. V. Ward, M. J. Butler, H. Sato, M. Seiki, J. J. Reynolds and G. Murphy (1995). "Intermolecular autolytic cleavage can contribute to the activation of progelatinase A by cell membranes." J Biol Chem **270**(51): 30479-30485.

Babine, R. E. and S. L. Bender (1997). "Molecular Recognition of Protein-Ligand Complexes: Applications to Drug Design." Chem Rev **97**(5): 1359-1472.

Bachinger, H. P. (1987). "The influence of peptidyl-prolyl cis-trans isomerase on the in vitro folding of type III collagen." J Biol Chem **262**(35): 17144-17148.

Bachinger, H. P., P. Bruckner, R. Timpl and J. Engel (1978). "The role of cis-trans isomerization of peptide bonds in the coil leads to and comes from triple helix conversion of collagen." Eur J Biochem **90**(3): 605-613.

Banh, A., J. Zhang, H. Cao, D. M. Bouley, S. Kwok, C. Kong, A. J. Giaccia, A. C. Koong and Q. T. Le (2011). "Tumor galectin-1 mediates tumor growth and metastasis through regulation of T-cell apoptosis." Cancer Res **71**(13): 4423-4431.

Bannikov, G. A., T. V. Karelina, I. E. Collier, B. L. Marmer and G. I. Goldberg (2002). "Substrate binding of gelatinase B induces its enzymatic activity in the presence of intact propeptide." J Biol Chem **277**(18): 16022-16027.

Banyai, L., H. Tordai and L. Patthy (1994). "The gelatin-binding site of human 72 kDa type IV collagenase (gelatinase A)." Biochem J **298 ( Pt 2)**: 403-407.

Bartlett, J. D., J. P. Simmer, J. Xue, H. C. Margolis and E. C. Moreno (1996). "Molecular cloning and mRNA tissue distribution of a novel matrix metalloproteinase isolated from porcine enamel organ." Gene **183**(1-2): 123-128.

Baselt, D. R., J. P. Revel and J. D. Baldeschwieler (1993). "Subfibrillar structure of type I collagen observed by atomic force microscopy." Biophys J **65**(6): 2644-2655.

Basset, P., J. P. Bellocq, C. Wolf, I. Stoll, P. Hutin, J. M. Limacher, O. L. Podhajcer, M. P. Chenard, M. C. Rio and P. Chambon (1990). "A novel metalloproteinase gene specifically expressed in stromal cells of breast carcinomas." Nature **348**(6303): 699-704.

Beck, K., V. C. Chan, N. Shenoy, A. Kirkpatrick, J. A. Ramshaw and B. Brodsky (2000). "Destabilization of osteogenesis imperfecta collagen-like model peptides correlates with the identity of the residue replacing glycine." Proc Natl Acad Sci U S A **97**(8): 4273-4278.

Becker, J. W., A. I. Marcy, L. L. Rokosz, M. G. Axel, J. J. Burbaum, P. M. Fitzgerald, P. M. Cameron, C. K. Esser, W. K. Hagmann, J. D. Hermes and et al. (1995). "Stromelysin-1: three-dimensional structure of the inhibited catalytic domain and of the C-truncated proenzyme." Protein Sci **4**(10): 1966-1976.

Belaaouaj, A., J. M. Shipley, D. K. Kobayashi, D. B. Zimonjic, N. Popescu, G. A. Silverman and S. D. Shapiro (1995). "Human macrophage metalloelastase. Genomic organization, chromosomal location, gene linkage, and tissue-specific expression." J Biol Chem **270**(24): 14568-14575.

Belkin, A. M., S. S. Akimov, L. S. Zaritskaya, B. I. Ratnikov, E. I. Deryugina and A. Y. Strongin (2001). "Matrix-dependent proteolysis of surface transglutaminase by membrane-type metalloproteinase regulates cancer cell adhesion and locomotion." J Biol Chem **276**(21): 18415-18422.

Bella, J., B. Brodsky and H. M. Berman (1995). "Hydration structure of a collagen peptide." Structure **3**(9): 893-906.

Bella, J., M. Eaton, B. Brodsky and H. M. Berman (1994). "Crystal and molecular structure of a collagen-like peptide at 1.9 A resolution." Science **266**(5182): 75-81.

Benbow, U., M. P. Schoenermark, T. I. Mitchell, J. L. Rutter, K. Shimokawa, H. Nagase and C. E. Brinckerhoff (1999). "A novel host/tumor cell interaction activates matrix metalloproteinase 1 and mediates invasion through type I collagen." J Biol Chem **274**(36): 25371-25378.

Bergt, C., G. Marsche, U. Panzenboeck, J. W. Heinecke, E. Malle and W. Sattler (2001). "Human neutrophils employ the myeloperoxidase/hydrogen peroxide/chloride system to oxidatively damage apolipoprotein A-I." Eur J Biochem **268**(12): 3523-3531.

Bertini, I., V. Calderone, M. Fragai, C. Luchinat, S. Mangani and B. Terni (2004). "Crystal structure of the catalytic domain of human matrix metalloproteinase 10." J Mol Biol **336**(3): 707-716.

Bertini, I., V. Calderone, M. Fragai, C. Luchinat, M. Maletta and K. J. Yeo (2006). "Snapshots of the reaction mechanism of matrix metalloproteinases." Angew Chem Int Ed Engl **45**(47): 7952-7955.

Bertini, I., V. Calderone, M. Fragai, R. Jaiswal, C. Luchinat, M. Melikian, E. Mylonas and D. I. Svergun (2008). "Evidence of reciprocal reorientation of the catalytic and hemopexin-like domains of full-length MMP-12." J Am Chem Soc **130**(22): 7011-7021.

Bertini, I., M. Fragai, C. Luchinat, M. Melikian, E. Mylonas, N. Sarti, D. I. Svergun, V. Maragliano, E. Molecular and H. Outstation (2009). "Interdomain Flexibility in Full-length Matrix Metalloproteinase-1 (MMP-1)" Journal of Biological Chemistry **284**: 12821-12828.

Bertini, I., M. Fragai, C. Luchinat, M. Melikian, M. Toccafondi, J. L. Lauer and G. B. Fields (2012). "Structural basis for matrix metalloproteinase 1-catalyzed collagenolysis." J Am Chem Soc **134**(4): 2100-2110.

Birkedal-Hansen, H., W. G. Moore, M. K. Bodden, L. J. Windsor, B. Birkedal-Hansen, A. DeCarlo and J. A. Engler (1993). "Matrix metalloproteinases: a review." Crit Rev Oral Biol Med **4**(2): 197-250.

Birkedal-Hansen, H., R. E. Taylor, A. S. Bhown, J. Katz, H. Y. Lin and B. R. Wells (1985). "Cleavage of bovine skin type III collagen by proteolytic enzymes. Relative resistance of the fibrillar form." J Biol Chem **260**(30): 16411-16417.

Birnboim, H. C. and J. Doly (1979). "A rapid alkaline extraction procedure for screening recombinant plasmid DNA." Nucleic Acids Res **7**(6): 1513-1523.

Bode, W., C. Fernandez-Catalan, H. Tschesche, F. Grams, H. Nagase and K. Maskos (1999). "Structural properties of matrix metalloproteinases." Cell Mol Life Sci **55**(4): 639-652.

Bode, W., F. X. Gomis-Ruth and W. Stockler (1993). "Astacins, serralysins, snake venom and matrix metalloproteinases exhibit identical zinc-binding environments (HEXXHXXGXXH and Met-turn) and topologies and should be grouped into a common family, the 'metzincins'." FEBS Lett **331**(1-2): 134-140.

Bode, W. and K. Maskos (2003). "Structural basis of the matrix metalloproteinases and their physiological inhibitors, the tissue inhibitors of metalloproteinases." Biological chemistry **384**: 863-872.

Boire, A., L. Covic, A. Agarwal, S. Jacques, S. Sherifi and A. Kuliopulos (2005). "PAR1 is a matrix metalloprotease-1 receptor that promotes invasion and tumorigenesis of breast cancer cells." Cell **120**(3): 303-313.

Bond, M., R. P. Fabunmi, A. H. Baker and A. C. Newby (1998). "Synergistic upregulation of metalloproteinase-9 by growth factors and inflammatory cytokines: an absolute requirement for transcription factor NF-kappa B." FEBS Lett **435**(1): 29-34.

Bonfanti, L., A. A. Mironov, Jr., J. A. Martinez-Menarguez, O. Martella, A. Fusella, M. Baldassarre, R. Buccione, H. J. Geuze, A. A. Mironov and A. Luini (1998). "Procollagen traverses the Golgi stack without leaving the lumen of cisternae: evidence for cisternal maturation." Cell **95**(7): 993-1003.

Borden, P., D. Solymar, A. Sucharczuk, B. Lindman, P. Cannon and R. A. Heller (1996). "Cytokine control of interstitial collagenase and collagenase-3 gene expression in human chondrocytes." J Biol Chem **271**(38): 23577-23581.

Bosshart, H., J. Humphrey, E. Deignan, J. Davidson, J. Drazba, L. Yuan, V. Oorschot, P. J. Peters and J. S. Bonifacino (1994). "The cytoplasmic domain mediates localization of furin to the trans-Golgi network en route to the endosomal/lysosomal system." J Cell Biol **126**(5): 1157-1172.

Bourboulia, D. and W. G. Stetler-Stevenson (2010). "Matrix metalloproteinases (MMPs) and tissue inhibitors of metalloproteinases (TIMPs): Positive and negative regulators in tumor cell adhesion." Semin Cancer Biol **20**(3): 161-168.

Bourguignon, L. Y., Z. Gunja-Smith, N. Iida, H. B. Zhu, L. J. Young, W. J. Muller and R. D. Cardiff (1998). "CD44v(3,8-10) is involved in cytoskeleton-mediated tumor cell migration and matrix metalloproteinase (MMP-9) association in metastatic breast cancer cells." J Cell Physiol **176**(1): 206-215.

Bourhis, J. M., N. Mariano, Y. Zhao, K. Harlos, J. Y. Exposito, E. Y. Jones, C. Moali, N. Aghajari and D. J. Hulmes (2012). "Structural basis of fibrillar collagen trimerization and related genetic disorders." Nat Struct Mol Biol.

Bowden, G. A., A. M. Paredes and G. Georgiou (1991). "Structure and morphology of protein inclusion bodies in Escherichia coli." Biotechnology (N Y) **9**(8): 725-730.

Brew, K. and H. Nagase (2010). "The tissue inhibitors of metalloproteinases (TIMPs): an ancient family with structural and functional diversity." Biochimica et biophysica acta **1803**: 55-71.

Broennimann, C., E. F. Eikenberry, B. Henrich, R. Horisberger, G. Huelsen, E. Pohl, B. Schmitt, C. Schulze-Briese, M. Suzuki, T. Tomizaki, H. Toyokawa and A. Wagner (2006). "The PILATUS 1M detector." J Synchrotron Radiat **13**(Pt 2): 120-130.

Brooks, P. C., S. Silletti, T. L. von Schalscha, M. Friedlander and D. A. Cheresh (1998). "Disruption of angiogenesis by PEX, a noncatalytic metalloproteinase fragment with integrin binding activity." Cell **92**(3): 391-400.

Brooks, P. C., S. Stromblad, L. C. Sanders, T. L. von Schalscha, R. T. Aimes, W. G. Stetler-Stevenson, J. P. Quigley and D. A. Cheresh (1996). "Localization of matrix metalloproteinase MMP-2 to the surface of invasive cells by interaction with integrin alpha v beta 3." Cell **85**(5): 683-693.

Bruckner, P. and D. J. Prockop (1981). "Proteolytic enzymes as probes for the triple-helical conformation of procollagen." Anal Biochem **110**(2): 360-368.

Butler, T. A., C. Zhu, R. A. Mueller, G. C. Fuller, W. J. Lemaire and J. F. Woessner, Jr. (1991). "Inhibition of ovulation in the perfused rat ovary by the synthetic collagenase inhibitor SC 44463." Biol Reprod **44**(6): 1183-1188.

Canty, E. G., Y. Lu, R. S. Meadows, M. K. Shaw, D. F. Holmes and K. E. Kadler (2004). "Coalignment of plasma membrane channels and protrusions (fibripositors) specifies the parallelism of tendon." J Cell Biol **165**(4): 553-563.

Carafoli, F., D. Bihan, S. Stathopoulos, A. D. Konitsiotis, M. Kvansakul, R. W. Farndale, B. Leitinger and E. Hohenester (2009). "Crystallographic insight into collagen recognition by discoidin domain receptor 2." Structure **17**(12): 1573-1581.

Caterina, J. J., Z. Skobe, J. Shi, Y. Ding, J. P. Simmer, H. Birkedal-Hansen and J. D. Bartlett (2002). "Enamelysin (matrix metalloproteinase 20)-deficient mice display an amelogenesis imperfecta phenotype." J Biol Chem **277**(51): 49598-49604.

Cawston, T. E., G. Murphy, E. Mercer, W. A. Galloway, B. L. Hazleman and J. J. Reynolds (1983). "The interaction of purified rabbit bone collagenase with purified rabbit bone metalloproteinase inhibitor." Biochem J **211**(2): 313-318.

Cha, H., E. Kopetzki, R. Huber, M. Lanzendorfer and H. Brandstetter (2002). "Structural basis of the adaptive molecular recognition by MMP9." J Mol Biol **320**(5): 1065-1079.

Chamberlin, M. and J. Ring (1973). "Characterization of T7-specific ribonucleic acid polymerase. 1. General properties of the enzymatic reaction and the template specificity of the enzyme." J Biol Chem **248**(6): 2235-2244.

Chandrasekaran, A. (2011). "Current trends in X-ray crystallography." from http://www.intechweb.org/books/show/title/current-trends-in-x-ray-crystallography.

Cheetham, G. M., D. Jeruzalmi and T. A. Steitz (1999). "Structural basis for initiation of transcription from an RNA polymerase-promoter complex." Nature **399**(6731): 80-83.

ChemAxon. (2010). MarvinSketch, Version 5.10.3.

Chen, L. C., M. E. Noelken and H. Nagase (1993). "Disruption of the cysteine-75 and zinc ion coordination is not sufficient to activate the precursor of human matrix metalloproteinase 3 (stromelysin 1)." Biochemistry **32**: 10289-10295.

Chen, L., Rydel, T. J., Gu, F., Dunaway, C. M., Pikul, S., Dunham, K. M., & Barnett, B. L. (1999). Crystal structure of the stromelysin catalytic domain at 2.0 A resolution: inhibitor-induced conformational changes. *J Mol Biol, 293*(3), 545-557. doi: 10.1006/jmbi.1999.3147

Chessler, S. D. and P. H. Byers (1993). "BiP binds type I procollagen pro alpha chains with mutations in the carboxyl-terminal propeptide synthesized by cells from patients with osteogenesis imperfecta." J Biol Chem **268**(24): 18226-18233.

Chetty, C., S. K. Vanamala, C. S. Gondi, D. H. Dinh, M. Gujrati and J. S. Rao (2012). "MMP-9 induces CD44 cleavage and CD44 mediated cell migration in glioblastoma xenograft cells." Cell Signal **24**(2): 549-559.

Chin, J. R., G. Murphy and Z. Werb (1985). "Stromelysin, a connective tissue-degrading metalloendopeptidase secreted by stimulated rabbit synovial fibroblasts in parallel with collagenase. Biosynthesis, isolation, characterization, and substrates." J Biol Chem **260**(22): 12367-12376.

Christensen, U. and L. Molgaard (1991). "Stopped-flow fluorescence kinetic studies of Glu-plasminogen. Conformational changes triggered by AH-site ligand binding." FEBS Lett **278**(2): 204-206.

Chung, L., N. Yoshida, J. L. Lauer-fields, G. B. Fields, R. Visse and H. Nagase (2004). "Collagenase unwinds triple-helical collagen prior to peptide bond hydrolysis." EMBO Journal **23**: 3020-3030.

Chung, L., K. Shimokawa, D. Dinakarpandian, F. Grams, G. B. Fields and H. Nagase (2000). "Identification of the (183)RWTNNFREY(191) region as a critical segment of matrix metalloproteinase 1 for the expression of collagenolytic activity." J Biol Chem **275**(38): 29610-29617.

Churg, A., S. Zhou and J. L. Wright (2012). "Series "matrix metalloproteinases in lung health and disease": Matrix metalloproteinases in COPD." Eur Respir J **39**(1): 197-209.

Clark, H. F., A. L. Gurney, et al. (2003). "The secreted protein discovery initiative (SPDI), a large-scale effort to identify novel human secreted and transmembrane proteins: a bioinformatics assessment." Genome Res **13**(10): 2265-2270.

Clark, I. M. and T. E. Cawston (1989). "Fragments of human fibroblast collagenase. Purification and characterization." Biochem J **263**(1): 201-206.

Clendeninn, N. J. and K. Appelt (2011). Matrix Metalloproteinase Inhibitors in Cancer Therapy, Humana Press.

Cole, A. A., S. Chubinskaya, B. Schumacher, K. Huch, G. Szabo, J. Yao, K. Mikecz, K. A. Hasty and K. E. Kuettner (1996). "Chondrocyte matrix metalloproteinase-8. Human articular chondrocytes express neutrophil collagenase." J Biol Chem **271**(18): 11023-11026.

Colige, A., I. Vandenberghe, M. Thiry, C. A. Lambert, J. Van Beeumen, S. W. Li, D. J. Prockop, C. M. Lapiere and B. V. Nusgens (2002). "Cloning and characterization of ADAMTS-14, a novel ADAMTS displaying high homology with ADAMTS-2 and ADAMTS-3." J Biol Chem **277**(8): 5756-5766.

Collier, I. E., W. Legant, B. Marmer, O. Lubman, S. Saffarian, T. Wakatsuki, E. Elson and G. I. Goldberg (2011). "Diffusion of MMPs on the Surface of Collagen Fibrils: The Mobile Cell Surface - Collagen Substratum Interface." PloS one **6**: e24029.

Crabbe, T., J. P. O'Connell, B. J. Smith and A. J. Docherty (1994). "Reciprocated matrix metalloproteinase activation: a process performed by interstitial collagenase and progelatinase A." Biochemistry **33**(48): 14419-14425.

Crawford, H. C., B. M. Fingleton, L. A. Rudolph-Owen, K. J. Goss, B. Rubinfeld, P. Polakis and L. M. Matrisian (1999). "The metalloproteinase matrilysin is a target of beta-catenin transactivation in intestinal tumors." Oncogene **18**(18): 2883-2891.

Csun.edu. (n.d.). "Novagen pET System " Retrieved 4 October, 2012, from http://www.csun.edu/~hcbio027/biotechnology/lec4a/petsys.html.

Cubellis, M. V., M. L. Nolli, G. Cassani and F. Blasi (1986). "Binding of single-chain prourokinase to the urokinase receptor of human U937 cells." J Biol Chem **261**(34): 15819-15822.

Cunningham, O., A. Andolfo, M. L. Santovito, L. Iuzzolino, F. Blasi and N. Sidenius (2003). "Dimerization controls the lipid raft partitioning of uPAR/CD87 and regulates its biological functions." EMBO J **22**(22): 5994-6003.

d'Ortho, M. P., H. Will, S. Atkinson, G. Butler, A. Messent, J. Gavrilovic, B. Smith, R. Timpl, L. Zardi and G. Murphy (1997). "Membrane-type matrix metalloproteinases 1 and 2 exhibit broad-spectrum proteolytic capacities comparable to many matrix metalloproteinases." Eur J Biochem **250**(3): 751-757.

Davis, J. M., B. A. Boswell and H. P. Bachinger (1989). "Thermal stability and folding of type IV procollagen and effect of peptidyl-prolyl cis-trans-isomerase on the folding of the triple helix." J Biol Chem **264**(15): 8956-8962.

de Coignac, A. B., G. Elson, Y. Delneste, G. Magistrelli, P. Jeannin, J. P. Aubry, O. Berthier, D. Schmitt, J. Y. Bonnefoy and J. F. Gauchat (2000). "Cloning of MMP-26. A novel matrilysin-like proteinase." Eur J Biochem **267**(11): 3323-3329.

De Souza, S. J., H. M. Pereira, S. Jacchieri and R. R. Brentani (1996). "Collagen/collagenase interaction: does the enzyme mimic the conformation of its own substrate?" FASEB J **10**(8): 927-930.

DeClerck, Y. A., T. D. Yean, H. S. Lu, J. Ting and K. E. Langley (1991). "Inhibition of autoproteolytic activation of interstitial procollagenase by recombinant metalloproteinase inhibitor MI/TIMP-2." J Biol Chem **266**(6): 3893-3899.

Decline, F. and P. Rousselle (2001). "Keratinocyte migration requires alpha2beta1 integrin-mediated interaction with the laminin 5 gamma2 chain." J Cell Sci **114**(Pt 4): 811-823.

Dectris® (2012). Pilatus3 Hybrid Pixel Detector Series Product Flyer.

del Mar Barbacid, M., P. Fernandez-Resa, J. M. Buesa, G. Marquez, M. Aracil, A. R. Quesadaand and E. Mira (1998). "Expression and purification of human stromelysin 1 and 3 from baculovirus-infected insect cells." Protein Expr Purif **13**(2): 243-250.

Delaglio, F., S. Grzesiek, G. W. Vuister, G. Zhu, J. Pfeifer and A. Bax (1995). "NMRPipe: a multidimensional spectral processing system based on UNIX pipes." J Biomol NMR **6**(3): 277-293.

Deryugina, E. I., B. Ratnikov, E. Monosov, T. I. Postnova, R. DiScipio, J. W. Smith and A. Y. Strongin (2001). "MT1-MMP initiates activation of pro-MMP-2 and integrin alphavbeta3 promotes maturation of MMP-2 in breast carcinoma cells." Exp Cell Res **263**(2): 209-223.

Desai, B., M. J. Rogers and M. A. Chellaiah (2007). "Mechanisms of osteopontin and CD44 as metastatic principles in prostate cancer cells." Mol Cancer **6**: 18.

Diaz, N. and D. Suarez (2012). "Alternative interdomain configurations of the full-length MMP-2 enzyme explored by molecular dynamics simulations." J Phys Chem B **116**(9): 2677-2686.

Doege, K. J. and J. H. Fessler (1986). "Folding of carboxyl domain and assembly of procollagen I." J Biol Chem **261**(19): 8924-8935.

Dombkowski, A. A. (2003). Disulfide by Design: a computational method for the rational design of disulfide bonds in proteins. *Bioinformatics, 19*(14), 1852-1853.

Drake, M. P., P. F. Davison, S. Bump and F. O. Schmitt (1966). "Action of proteolytic enzymes on tropocollagen and insoluble collagen." Biochemistry **5**(1): 301-312.

Dufour, A., S. Zucker, N. S. Sampson, C. Kuscu and J. Cao (2010). "Role of matrix metalloproteinase-9 dimers in cell migration: design of inhibitory peptides." J Biol Chem **285**(46): 35944-35956.

Dumin, J. A., S. K. Dickeson, T. P. Stricker, M. Bhattacharyya-Pakrasi, J. D. Roby, S. A. Santoro and W. C. Parks (2001). "Pro-collagenase-1 (matrix metalloproteinase-1) binds the alpha(2)beta(1) integrin upon release from keratinocytes migrating on type I collagen." J Biol Chem **276**(31): 29368-29374.

Eck, S. M., A. L. Cote, W. D. Winkelman and C. E. Brinckerhoff (2009). "CXCR4 and matrix metalloproteinase-1 are elevated in breast carcinoma-associated fibroblasts and in normal

mammary fibroblasts exposed to factors secreted by breast cancer cells." Mol Cancer Res **7**(7): 1033-1044.

Elkins, P. A., Y. S. Ho, W. W. Smith, C. A. Janson, K. J. D'Alessio, M. S. McQueney, M. D. Cummings and A. M. Romanic (2002). "Structure of the C-terminally truncated human ProMMP9, a gelatin-binding matrix metalloproteinase." Acta Crystallogr D Biol Crystallogr **58**(Pt 7): 1182-1192.

Ellerbroek, S. M., D. A. Fishman, A. S. Kearns, L. M. Bafetti and M. S. Stack (1999). "Ovarian carcinoma regulation of matrix metalloproteinase-2 and membrane type 1 matrix metalloproteinase through beta1 integrin." Cancer Res **59**(7): 1635-1641.

Ellis, V., S. A. Whawell, F. Werner and J. J. Deadman (1999). "Assembly of urokinase receptor-mediated plasminogen activation complexes involves direct, non-active-site interactions between urokinase and plasminogen." Biochemistry **38**(2): 651-659.

Emsley, J., C. G. Knight, R. W. Farndale, M. J. Barnes and R. C. Liddington (2000). "Structural basis of collagen recognition by integrin alpha2beta1." Cell **101**(1): 47-56.

Endo, K., T. Takino, H. Miyamori, H. Kinsen, T. Yoshizaki, M. Furukawa and H. Sato (2003). "Cleavage of syndecan-1 by membrane type matrix metalloproteinase-1 stimulates cell migration." J Biol Chem **278**(42): 40764-40770.

Enghild, J. J., G. Salvesen, K. Brew and H. Nagase (1989). "Interaction of human rheumatoid synovial collagenase (matrix metalloproteinase 1) and stromelysin (matrix metalloproteinase 3) with human alpha 2-macroglobulin and chicken ovostatin. Binding kinetics and identification of matrix metalloproteinase cleavage sites." J Biol Chem **264**(15): 8779-8785.

Evanson, J. M., J. J. Jeffrey and S. M. Krane (1968). "Studies on collagenase from rheumatoid synovium in tissue culture." J Clin Invest **47**(12): 2639-2651.

Eyre, D. R., M. A. Weis and J. J. Wu (2008). "Advances in collagen cross-link analysis." Methods **45**(1): 65-74.

Faber, H. R., C. R. Groom, H. M. Baker, W. T. Morgan, A. Smith and E. N. Baker (1995). "1.8 A crystal structure of the C-terminal domain of rabbit serum haemopexin." Structure **3**(6): 551-559.

Feigin, L. A. and D. I. Svergun (1987). Structure Analysis by Small-Angle X-Ray and Neutron Scattering, Plenum Press/Springer.

Feng, Y., Likos, J. J., Zhu, L., Woodward, H., Munie, G., McDonald, J. J., Stallings, W. C. (2002). Solution structure and backbone dynamics of the catalytic domain of matrix metalloproteinase-2 complexed with a hydroxamic acid inhibitor. *Biochim Biophys Acta, 1598*(1-2), 10-23.

Fernandes, R. J., S. Hirohata, J. M. Engle, A. Colige, D. H. Cohn, D. R. Eyre and S. S. Apte (2001). "Procollagen II amino propeptide processing by ADAMTS-3. Insights on dermatosparaxis." J Biol Chem **276**(34): 31502-31509.

Fernandez, J. M. and J. P. Hoeffler (1998). Gene Expression Systems: Using Nature for the Art of Expression. San Diego, Academic Press.

Fernandez-catalan, C., W. Bode, R. Huber, D. Turk, J. J. Calvete, A. Lichte, H. Tschesche and K. Maskos (1998). "Crystal structure of the complex formed by the membrane type 1-matrix metalloproteinase with the tissue inhibitor of metalloproteinases-2, the soluble progelatinase A receptor." EMBO Journal **17**: 5238-5248.

Fields, G. B. (1991). "A model for interstitial collagen catabolism by mammalian collagenases." J Theor Biol **153**(4): 585-602.

Fields, G. B. (2010). "Synthesis and biological applications of collagen-model triple-helical peptides." Org Biomol Chem **8**(6): 1237-1258.

Fields, G. B. and D. J. Prockop (1996). "Perspectives on the synthesis and application of triple-helical, collagen-model peptides." Biopolymers **40**(4): 345-357.

Fiori, S., B. Sacca and L. Moroder (2002). "Structural properties of a collagenous heterotrimer that mimics the collagenase cleavage site of collagen type I." J Mol Biol **319**(5): 1235-1242.

Formstone, C. J., P. J. Byrd, H. J. Ambrose, J. H. Riley, D. Hernandez, C. M. McConville and A. M. Taylor (1993). "The order and orientation of a cluster of metalloproteinase genes, stromelysin 2, collagenase, and stromelysin, together with D11S385, on chromosome 11q22-q23." Genomics **16**(1): 289-291.

Fosang, A. J., P. J. Neame, K. Last, T. E. Hardingham, G. Murphy and J. A. Hamilton (1992). "The interglobular domain of cartilage aggrecan is cleaved by PUMP, gelatinases, and cathepsin B." J Biol Chem **267**(27): 19470-19474.

Franke, D. and D. I. Svergun (2009). "DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering." Journal of Applied Crystallography **42**(2): 342-346.

Freije, J. M., I. Diez-Itza, M. Balbin, L. M. Sanchez, R. Blasco, J. Tolivia and C. Lopez-Otin (1994). "Molecular cloning and expression of collagenase-3, a novel human matrix metalloproteinase produced by breast carcinomas." J Biol Chem **269**(24): 16766-16773.

Freimark, B. D., W. S. Feeser and S. A. Rosenfeld (1994). "Multiple sites of the propeptide region of human stromelysin-1 are required for maintaining a latent form of the enzyme." J Biol Chem **269**(43): 26982-26987.

Fridman, R., M. Toth, D. Pena and S. Mobashery (1995). "Activation of progelatinase B (MMP-9) by gelatinase A (MMP-2)." Cancer Res **55**(12): 2548-2555.

Fu, X., S. Y. Kassim, W. C. Parks and J. W. Heinecke (2001). "Hypochlorous acid oxygenates the cysteine switch domain of pro-matrilysin (MMP-7). A mechanism for matrix metalloproteinase activation and atherosclerotic plaque rupture by myeloperoxidase." J Biol Chem **276**(44): 41279-41287.

Galazka, G., L. J. Windsor, H. Birkedal-Hansen and J. A. Engler (1996). "APMA (4-aminophenylmercuric acetate) activation of stromelysin-1 involves protein interactions in addition to those with cysteine-75 in the propeptide." Biochemistry **35**(34): 11221-11227.

Galli, A., G. Svegliati-Baroni, E. Ceni, S. Milani, F. Ridolfi, R. Salzano, M. Tarocchi, C. Grappone, G. Pellegrini, A. Benedetti, C. Surrenti and A. Casini (2005). "Oxidative stress stimulates proliferation and invasiveness of hepatic stellate cells via a MMP2-mediated mechanism." Hepatology **41**(5): 1074-1084.

Galloway, W. A., G. Murphy, J. D. Sandy, J. Gavrilovic, T. E. Cawston and J. J. Reynolds (1983). "Purification and characterization of a rabbit bone metalloproteinase that degrades proteoglycan and other connective-tissue components." Biochem J **209**(3): 741-752.

Galvez, B. G., S. Matias-Roman, J. P. Albar, F. Sanchez-Madrid and A. G. Arroyo (2001). "Membrane type 1-matrix metalloproteinase is activated during migration of human endothelial cells and modulates endothelial motility and matrix remodeling." J Biol Chem **276**(40): 37491-37500.

Galvez, B. G., S. Matias-Roman, M. Yanez-Mo, F. Sanchez-Madrid and A. G. Arroyo (2002). "ECM regulates MT1-MMP localization with beta1 or alphavbeta3 integrins at distinct cell compartments modulating its internalization and activity on human endothelial cells." J Cell Biol **159**(3): 509-521.

Ganz, P. R., D. Dupuis, A. K. Dudani and S. Hashemi (1991). "Characterization of plasminogen binding to human capillary and arterial endothelial cells." Biochem Cell Biol **69**(7): 442-448.

Garnero, P., O. Borel, I. Byrjalsen, M. Ferreras, F. H. Drake, M. S. McQueney, N. T. Foged, P. D. Delmas and J. M. Delaisse (1998). "The collagenolytic activity of cathepsin K is unique among mammalian proteinases." J Biol Chem **273**(48): 32347-32352.

Gasteiger, E., A. Gattiker, C. Hoogland, I. Ivanyi, R. D. Appel and A. Bairoch (2003). "ExPASy: The proteomics server for in-depth protein knowledge and analysis." Nucleic Acids Res **31**(13): 3784-3788.

GE Healthcare (2008). Biacore - Sensor Surface Handbook.

Giambernardi, T. A., A. Y. Sakaguchi, J. Gluhak, D. Pavlin, D. A. Troyer, G. Das, U. Rodeck and R. J. Klebe (2001). "Neutrophil collagenase (MMP-8) is expressed during early development in neural crest cells as well as in adult melanoma cells." Matrix Biol **20**(8): 577-587.

Gilles, M. A., A. Q. Hudson and C. L. Borders, Jr. (1990). "Stability of water-soluble carbodiimides in aqueous solution." Anal Biochem **184**(2): 244-248.

Gingras, D., N. Bousquet-Gagnon, S. Langlois, M. P. Lachambre, B. Annabi and R. Beliveau (2001). "Activation of the extracellular signal-regulated protein kinase (ERK) cascade by membrane-type-1 matrix metalloproteinase (MT1-MMP)." FEBS Lett **507**(2): 231-236.

Gioia, M., S. Monaco, G. F. Fasciglione, A. Coletti, A. Modesti, S. Marini and M. Coletta (2007). "Characterization of the mechanisms by which gelatinase A, neutrophil collagenase, and membrane-type metalloproteinase MMP-14 recognize collagen I and enzymatically process the two alpha-chains." J Mol Biol **368**(4): 1101-1113.

Glatter, O. (1977). "A new method for the evaluation of small-angle scattering data." Journal of Applied Crystallography **10**(5): 415-421.

Gohlke, U., F. X. Gomis-Ruth, T. Crabbe, G. Murphy, A. J. Docherty and W. Bode (1996). "The C-terminal (haemopexin-like) domain structure of human gelatinase A (MMP2): structural implications for its function." FEBS Lett **378**(2): 126-130.

Goldberg, G. I., S. M. Wilhelm, A. Kronberger, E. A. Bauer, G. A. Grant and A. Z. Eisen (1986). "Human fibroblast collagenase. Complete primary structure and homology to an oncogene transformation-induced rat protein." J Biol Chem **261**(14): 6600-6605.

Golubkov, V. S., A. V. Chekanov, S. A. Shiryaev, A. E. Aleshin, B. I. Ratnikov, K. Gawlik, I. Radichev, K. Motamedchaboki, J. W. Smith and A. Y. Strongin (2007). "Proteolysis of the membrane type-1 matrix metalloproteinase prodomain: implications for a two-step proteolytic processing and activation." J Biol Chem **282**(50): 36283-36291.

Golubkov, V. S., P. Cieplak, A. V. Chekanov, B. I. Ratnikov, A. E. Aleshin, N. V. Golubkova, T. I. Postnova, I. a. Radichev, D. V. Rozanov, W. Zhu, K. Motamedchaboki and A. Y. Strongin (2010). "Internal cleavages of the autoinhibitory prodomain are required for membrane type-1 matrix metalloproteinase activation while furin cleavage alone generates inactive proteinase." The Journal of biological chemistry **285**: 27726 -27736.

Gomis-Ruth, F. X., U. Gohlke, M. Betz, V. Knauper, G. Murphy, C. Lopez-Otin and W. Bode (1996). "The helping hand of collagenase-3 (MMP-13): 2.7 A crystal structure of its C-terminal haemopexin-like domain." J Mol Biol **264**(3): 556-566.

Gomis-Ruth, F. X., K. Maskos, M. Betz, A. Bergner, R. Huber, K. Suzuki, N. Yoshida, H. Nagase, K. Brew, G. P. Bourenkov, H. Bartunik and W. Bode (1997). "Mechanism of inhibition of the human matrix metalloproteinase stromelysin-1 by TIMP-1." Nature **389**(6646): 77-81.

Grant, G. A., A. Z. Eisen, B. L. Marmer, W. T. Roswit and G. I. Goldberg (1987). "The activation of human skin fibroblast procollagenase. Sequence identification of the major conversion products." J Biol Chem **262**(12): 5886-5889.

Grantham, R., C. Gautier, M. Gouy, R. Mercier and A. Pave (1980). "Codon catalog usage and the genome hypothesis." Nucleic Acids Res **8**(1): r49-r62.

Green, K. A., K. Almholt, M. Ploug, B. Rono, F. J. Castellino, M. Johnsen, T. H. Bugge, J. Romer and L. R. Lund (2008). "Profibrinolytic effects of metalloproteinases during skin wound healing in the absence of plasminogen." J Invest Dermatol **128**(8): 2092-2101.

Gross, J. and C. M. Lapiere (1962). "Collagenolytic activity in amphibian tissues: a tissue culture assay." Proc Natl Acad Sci U S A **48**: 1014-1022.

Gross, J. and Y. Nagai (1965). "Specific degradation of the collagen molecule by tadpole collagenolytic enzyme." Proc Natl Acad Sci U S A **54**(4): 1197-1204.

Gruber, H. E., J. A. Ingram, G. L. Hoelscher, N. Zinchenko, H. J. Norton and E. N. Hanley, Jr. (2009). "Matrix metalloproteinase 28, a novel matrix metalloproteinase, is constitutively expressed in human intervertebral disc tissue and is present in matrix of more degenerated discs." Arthritis Res Ther **11**(6): R184.

Gu, Z., M. Kaul, B. Yan, S. J. Kridel, J. Cui, A. Strongin, J. W. Smith, R. C. Liddington and S. A. Lipton (2002). "S-nitrosylation of matrix metalloproteinases: signaling pathway to neuronal cell death." Science **297**(5584): 1186-1190.

Gururajan, R., J. Grenet, J. M. Lahti and V. J. Kidd (1998). "Isolation and characterization of two novel metalloproteinase genes linked to the Cdc2L locus on human chromosome 1p36.3." Genomics **52**(1): 101-106.

Gururajan, R., J. M. Lahti, J. Grenet, J. Easton, I. Gruber, P. F. Ambros and V. J. Kidd (1998). "Duplication of a genomic region containing the Cdc2L1-2 and MMP21-22 genes on human chromosome 1p36.3 and their linkage to D1Z2." Genome Res **8**(9): 929-939.

Gutsmann, T., G. E. Fantner, J. H. Kindt, M. Venturoni, S. Danielsen and P. K. Hansma (2004). "Force spectroscopy of collagen fibers to investigate their mechanical properties and structural organization." Biophys J **86**(5): 3186-3193.

Hajjar, K. A., P. C. Harpel, E. A. Jaffe and R. L. Nachman (1986). "Binding of plasminogen to cultured human endothelial cells." J Biol Chem **261**(25): 11656-11662.

Han, S., E. Makareeva, N. V. Kuznetsova, A. M. DeRidder, M. B. Sutter, W. Losert, C. L. Phillips, R. Visse, H. Nagase and S. Leikin (2010). "Molecular mechanism of type I collagen homotrimer resistance to mammalian collagenases." J Biol Chem **285**(29): 22276-22281.

Hanemaaijer, R., P. Koolwijk, L. le Clercq, W. J. de Vree and V. W. van Hinsbergh (1993). "Regulation of matrix metalloproteinase expression in human vein and microvascular endothelial cells. Effects of tumour necrosis factor alpha, interleukin 1 and phorbol ester." Biochem J **296 ( Pt 3)**: 803-809.

Hardingham, T. E. and A. J. Fosang (1992). "Proteoglycans: many forms and many functions." FASEB J **6**(3): 861-870.

Harris, R., E. Becker, S. Cabral de Menezes, R. Goodfellow and P. Granger (2001). "NMR nomenclature. Nuclear spin properties and conventions for chemical shifts (IUPAC Recommendations 2001)." Pue Appl. Chem. **73**(11): 1795-1818.

Harwood, R., M. E. Grant and D. S. Jackson (1974). "Collagen biosynthesis. Characterization of subcellular fractions from embyonic chick fibroblasts and the intracellular localization of protocollagen prolyl and protocollagen lysyl hydroxylases." Biochem J **144**(1): 123-130.

Harwood, R., M. E. Grant and D. S. Jackson (1975). "Studies on the glycosylation of hydroxylysine residues during collagen biosynthesis and the subcellular localization of collagen galactosyltransferase and collagen glucosyltransferase in tendon and cartilage cells." Biochem J **152**(2): 291-302.

Hasty, K. A., J. J. Jeffrey, M. S. Hibbs and H. G. Welgus (1987). "The collagen substrate specificity of human neutrophil collagenase." J Biol Chem **262**(21): 10048-10052.

Hasty, K. A., T. F. Pourmotabbed, G. I. Goldberg, J. P. Thompson, D. G. Spinella, R. M. Stevens and C. L. Mainardi (1990). "Human neutrophil collagenase. A distinct gene product with homology to other matrix metalloproteinases." J Biol Chem **265**(20): 11421-11424.

He, C. S., S. M. Wilhelm, A. P. Pentland, B. L. Marmer, G. A. Grant, A. Z. Eisen and G. I. Goldberg (1989). "Tissue cooperation in a proteolytic cascade activating human interstitial collagenase." Proc Natl Acad Sci U S A **86**(8): 2632-2636.

Hedstrom, L. (2002). "Serine protease mechanism and specificity." Chem Rev **102**(12): 4501-4524.

Hembrough, T. A., L. Li and S. L. Gonias (1996). "Cell-surface cytokeratin 8 is the major plasminogen receptor on breast cancer cells and is required for the accelerated activation of cell-associated plasminogen by tissue-type plasminogen activator." J Biol Chem **271**(41): 25684-25691.

Henrich, S., A. Cameron, G. P. Bourenkov, R. Kiefersauer, R. Huber, I. Lindberg, W. Bode and M. E. Than (2003). "The crystal structure of the proprotein processing proteinase furin explains its stringent specificity." Nat Struct Biol **10**(7): 520-526.

Hibbs, M. S., K. A. Hasty, J. M. Seyer, A. H. Kang and C. L. Mainardi (1985). "Biochemical and immunological characterization of the secreted forms of human neutrophil gelatinase." J Biol Chem **260**(4): 2493-2500.

Hirose, T., C. Patterson, T. Pourmotabbed, C. L. Mainardi and K. A. Hasty (1993). "Structure-function relationship of human neutrophil collagenase: identification of regions responsible for substrate specificity and general proteinase activity." Proc Natl Acad Sci U S A **90**(7): 2569-2573.

Hofmann, U. B., J. R. Westphal, A. A. Van Kraats, D. J. Ruiter and G. N. Van Muijen (2000). "Expression of integrin alpha(v)beta(3) correlates with activation of membrane-type matrix metalloproteinase-1 (MT1-MMP) and matrix metalloproteinase-2 (MMP-2) in human melanoma cells in vitro and in vivo." Int J Cancer **87**(1): 12-19.

Horwitz, A. L., A. J. Hance and R. G. Crystal (1977). "Granulocyte collagenase: selective digestion of type I relative to type III collagen." Proc Natl Acad Sci U S A **74**(3): 897-901.

Housley, T. J., A. P. Baumann, I. D. Braun, G. Davis, P. K. Seperack and S. M. Wilhelm (1993). "Recombinant Chinese hamster ovary cell matrix metalloprotease-3 (MMP-3, stromelysin-1). Role of calcium in promatrix metalloprotease-3 (pro-MMP-3, prostromelysin-1) activation and thermostability of the low mass catalytic domain of MMP-3." J Biol Chem **268**(6): 4481-4487.

Huhtala, P., L. T. Chow and K. Tryggvason (1990). "Structure of the human type IV collagenase gene." J Biol Chem **265**(19): 11077-11082.

Huntington, J. T., J. M. Shields, C. J. Der, C. A. Wyatt, U. Benbow, C. L. Slingluff, Jr. and C. E. Brinckerhoff (2004). "Overexpression of collagenase 1 (MMP-1) is mediated by the ERK pathway in invasive melanoma cells: role of BRAF mutation and fibroblast growth factor signaling." J Biol Chem **279**(32): 33168-33176.

Hurst, D. R., M. A. Schwartz, M. A. Ghaffari, Y. Jin, H. Tschesche, G. B. Fields and Q. X. Sang (2004). "Catalytic- and ecto-domains of membrane type 1-matrix metalloproteinase have similar inhibition profiles but distinct endopeptidase activities." Biochem J **377**(Pt 3): 775-779.

Huxley-Jones, J., T. K. Clarke, C. Beck, G. Toubaris, D. L. Robertson and R. P. Boot-Handford (2007). "The evolution of the vertebrate metzincins; insights from Ciona intestinalis and Danio rerio." BMC Evol Biol **7**: 63.

Illman, S. A., J. Keski-Oja and J. Lohi (2001). "Promoter characterization of the human and mouse epilysin (MMP-28) genes." Gene **275**(1): 185-194.

Inada, M., Y. Wang, M. H. Byrne, M. U. Rahman, C. Miyaura, C. Lopez-Otin and S. M. Krane (2004). "Critical roles for collagenase-3 (Mmp13) in development of growth plate cartilage and in endochondral ossification." Proc Natl Acad Sci U S A **101**(49): 17192-17197.

Ito, A. and H. Nagase (1988). "Evidence that human rheumatoid synovial matrix metalloproteinase 3 is an endogenous activator of procollagenase." Arch Biochem Biophys **267**(1): 211-216.

Itoh, Y., S. Binner and H. Nagase (1995). "Steps involved in activation of the complex of pro-matrix metalloproteinase 2 (progelatinase A) and tissue inhibitor of metalloproteinases (TIMP)-2 by 4-aminophenylmercuric acetate." Biochem J **308 ( Pt 2)**: 645-651.

Itoh, Y., M. Kajita, H. Kinoh, H. Mori, A. Okada and M. Seiki (1999). "Membrane type 4 matrix metalloproteinase (MT4-MMP, MMP-17) is a glycosylphosphatidylinositol-anchored proteinase." J Biol Chem **274**(48): 34260-34266.

Itoh, Y., A. Takamura, N. Ito, Y. Maru, H. Sato, N. Suenaga, T. Aoki and M. Seiki (2001). "Homophilic complex formation of MT1-MMP facilitates proMMP-2 activation on the cell surface and promotes tumor cell invasion." EMBO J **20**(17): 4782-4793.

Iyer, S., R. Visse, H. Nagase and K. R. Acharya (2006). "Crystal Structure of an Active Form of Human MMP-1." Structure: 78-88.

Iyer, S., S. Wei, K. Brew, K. R. Acharya, C. Down, U. Kingdom and B. Raton (2007). "Crystal Structure of the Catalytic Domain of Matrix Metalloproteinase-1 in Complex with the Inhibitory Domain of Tissue Inhibitor of Metalloproteinase-1 *." Journal of Biological Chemistry **282**: 364 -371.

Johansson, N., K. Airola, R. Grenman, A. L. Kariniemi, U. Saarialho-Kere and V. M. Kahari (1997). "Expression of collagenase-3 (matrix metalloproteinase-13) in squamous cell carcinomas of the head and neck." Am J Pathol **151**(2): 499-508.

Jozic, D., G. Bourenkov, N. H. Lim, R. Visse, H. Nagase, W. Bode and K. Maskos (2005). "X-ray structure of human proMMP-1: new insights into procollagenase activation and collagen binding." J Biol Chem **280**(10): 9578-9585.

Juncker-Jensen, A. and L. R. Lund (2011). "Phenotypic overlap between MMP-13 and the plasminogen activation system during wound healing in mice." PloS one **6**(2): e16954.

Kadler, K. E., Y. Hojima and D. J. Prockop (1987). "Assembly of collagen fibrils de novo by cleavage of the type I pC-collagen with procollagen C-proteinase. Assay of critical concentration demonstrates that collagen self-assembly is a classical example of an entropy-driven process." J Biol Chem **262**(32): 15696-15701.

Kafienah, W., D. Bromme, D. J. Buttle, L. J. Croucher and A. P. Hollander (1998). "Human cathepsin K cleaves native type I and II collagens at the N-terminal end of the triple helix." Biochem J **331 ( Pt 3)**: 727-732.

Kajita, M., Y. Itoh, T. Chiba, H. Mori, A. Okada, H. Kinoh and M. Seiki (2001). "Membrane-type 1 matrix metalloproteinase cleaves CD44 and promotes cell migration." J Cell Biol **153**(5): 893-904.

Kane, J. F. (1995). "Effects of rare codon clusters on high-level expression of heterologous proteins in Escherichia coli." Curr Opin Biotechnol **6**(5): 494-500.

Kellokumpu, S., R. Sormunen, J. Heikkinen and R. Myllyla (1994). "Lysyl hydroxylase, a collagen processing enzyme, exemplifies a novel class of luminally-oriented peripheral membrane proteins in the endoplasmic reticulum." J Biol Chem **269**(48): 30524-30529.

Kessler, E., K. Takahara, L. Biniaminov, M. Brusel and D. S. Greenspan (1996). "Bone morphogenetic protein-1: the type I procollagen C-proteinase." Science **271**(5247): 360-362.

Khoshnoodi, J., J. P. Cartailler, K. Alvares, A. Veis and B. G. Hudson (2006). "Molecular recognition in the assembly of collagens: terminal noncollagenous domains are key recognition modules in the formation of triple helical protomers." J Biol Chem **281**(50): 38117-38121.

Kim, J. W., J. P. Simmer, T. C. Hart, P. S. Hart, M. D. Ramaswami, J. D. Bartlett and J. C. Hu (2005). "MMP-20 mutation in autosomal recessive pigmented hypomaturation amelogenesis imperfecta." J Med Genet **42**(3): 271-275.

Klawitter, M., L. Quero, A. Bertolo, M. Mehr, J. Stoyanov, A. G. Nerlich, J. Klasen, N. Aebli, N. Boos and K. Wuertz (2011). "Human MMP28 expression is unresponsive to inflammatory stimuli and does not correlate to the grade of intervertebral disc degeneration." J Negat Results Biomed **10**: 9.

Kleifeld, O., A. Doucet, U. auf dem Keller, A. Prudova, O. Schilling, R. K. Kainthan, A. E. Starr, L. J. Foster, J. N. Kizhakkedathu and C. M. Overall (2010). "Isotopic labeling of terminal amines in complex samples identifies protein N-termini and protease cleavage products." Nat Biotechnol **28**(3): 281-288.

Knauper, V., L. Bailey, J. R. Worley, P. Soloway, M. L. Patterson and G. Murphy (2002). "Cellular activation of proMMP-13 by MT1-MMP depends on the C-terminal domain of MMP-13." FEBS Lett **532**(1-2): 127-130.

Knauper, V., S. Cowell, B. Smith, C. Lopez-Otin, M. O'Shea, H. Morris, L. Zardi and G. Murphy (1997). "The role of the C-terminal domain of human collagenase-3 (MMP-13) in the activation of

procollagenase-3, substrate specificity, and tissue inhibitor of metalloproteinase interaction." J Biol Chem **272**(12): 7608-7616.

Knauper, V., S. Kramer, H. Reinke and H. Tschesche (1990). "Characterization and activation of procollagenase from human polymorphonuclear leucocytes. N-terminal sequence determination of the proenzyme and various proteolytically activated forms." Eur J Biochem **189**(2): 295-300.

Knauper, V., C. Lopez-Otin, B. Smith, G. Knight and G. Murphy (1996). "Biochemical characterization of human collagenase-3." J Biol Chem **271**(3): 1544-1550.

Knauper, V., G. Murphy and H. Tschesche (1996). "Activation of human neutrophil procollagenase by stromelysin 2." Eur J Biochem **235**(1-2): 187-191.

Knauper, V., S. M. Wilhelm, P. K. Seperack, Y. A. DeClerck, K. E. Langley, A. Osthues and H. Tschesche (1993). "Direct activation of human neutrophil procollagenase by recombinant stromelysin." Biochem J **295 ( Pt 2)**: 581-586.

Knauper, V., H. Will, C. Lopez-Otin, B. Smith, S. J. Atkinson, H. Stanton, R. M. Hembry and G. Murphy (1996). "Cellular mechanisms for human procollagenase-3 (MMP-13) activation. Evidence that MT1-MMP (MMP-14) and gelatinase a (MMP-2) are able to generate active enzyme." J Biol Chem **271**(29): 17124-17131.

Koide, T. (2007). "Designed triple-helical peptides as tools for collagen biochemistry and matrix engineering." Philos Trans R Soc Lond B Biol Sci **362**(1484): 1281-1291.

Kojima, S., Y. Itoh, S. Matsumoto, Y. Masuho and M. Seiki (2000). "Membrane-type 6 matrix metalloproteinase (MT6-MMP, MMP-25) is the second glycosyl-phosphatidyl inositol (GPI)-anchored MMP." FEBS Lett **480**(2-3): 142-146.

Kolb, C., S. Mauch, H. H. Peter, U. Krawinkel and R. Sedlacek (1997). "The matrix metalloproteinase RASI-1 is expressed in synovial blood vessels of a rheumatoid arthritis patient." Immunol Lett **57**(1-3): 83-88.

Konarev, P. V., M. V. Petoukhov, V. V. Volkov and D. I. Svergun (2006). "ATSAS 2.1, a program package for small-angle scattering data analysis." Journal of Applied Crystallography **39**(2): 277-286.

Konarev, P. V., V. V. Volkov, A. V. Sokolova, M. H. J. Koch and D. I. Svergun (2003). "PRIMUS: a Windows PC-based system for small-angle scattering data analysis." Journal of Applied Crystallography **36**(5): 1277-1282.

Konttinen, Y. T., M. Ainola, H. Valleala, J. Ma, H. Ida, J. Mandelin, R. W. Kinne, S. Santavirta, T. Sorsa, C. Lopez-Otin and M. Takagi (1999). "Analysis of 16 different matrix metalloproteinases (MMP-1 to MMP-20) in the synovial membrane: different profiles in trauma and rheumatoid arthritis." Ann Rheum Dis **58**(11): 691-697.

Konttinen, Y. T., A. Ceponis, M. Takagi, M. Ainola, T. Sorsa, M. Sutinen, T. Salo, J. Ma, S. Santavirta and M. Seiki (1998). "New collagenolytic enzymes/cascade identified at the pannus-hard tissue junction in rheumatoid arthritis: destruction from above." Matrix Biol **17**(8-9): 585-601.

Kozin, M. B. and D. I. Svergun (2001). "Automated matching of high- and low-resolution structural models." Journal of Applied Crystallography **34**(1): 33-41.

Kryczka, J., M. Stasiak, L. Dziki, M. Mik, A. Dziki and C. Cierniewski (2012). "Matrix metalloproteinase-2 cleavage of the beta1 integrin ectodomain facilitates colon cancer cell motility." J Biol Chem **287**(43): 36556-36566.

Lauer-Fields, J. L. and G. B. Fields (2002). "Triple-helical peptide analysis of collagenolytic protease activity." Biol Chem **383**(7-8): 1095-1105.

Lauer-Fields, J. L., D. Juska and G. B. Fields (2002). "Matrix metalloproteinases and collagen catabolism." Biopolymers **66**: 19-32.

Lauer-Fields, J. L., M. J. Chalmers, S. A. Busby, D. Minond, P. R. Griffin and G. B. Fields (2009). "Identification of specific hemopexin-like domain residues that facilitate matrix metalloproteinase collagenolytic activity." J Biol Chem **284**(36): 24017-24024.

Lees, J. F., M. Tasab and N. J. Bulleid (1997). "Identification of the molecular recognition sequence which determines the type-specific assembly of procollagen." EMBO J **16**(5): 908-916.

Lehti, K., J. Lohi, H. Valtanen and J. Keski-Oja (1998). "Proteolytic processing of membrane-type-1 matrix metalloproteinase is associated with gelatinase A activation at the cell surface." Biochem J **334 ( Pt 2)**: 345-353.

Lehti, K., H. Valtanen, S. A. Wickstrom, J. Lohi and J. Keski-Oja (2000). "Regulation of membrane-type-1 matrix metalloproteinase activity by its cytoplasmic domain." J Biol Chem **275**(20): 15006-15013.

Leikina, E., M. V. Mertts, N. Kuznetsova and S. Leikin (2002). "Type I collagen is thermally unstable at body temperature." Proceedings of the National Academy of Sciences of the United States of America **99**: 1314-1318.

Leung, M. K., L. I. Fessler, D. B. Greenberg and J. H. Fessler (1979). "Separate amino and carboxyl procollagen peptidases in chick embryo tendon." J Biol Chem **254**(1): 224-232.

Li, J., P. Brick, M. C. O'Hare, T. Skarzynski, L. F. Lloyd, V. a. Curry, I. M. Clark, H. F. Bigg, B. L. Hazleman and T. E. Cawston (1995). "Structure of full-length porcine synovial collagenase reveals a C-terminal domain containing a calcium-linked, four-bladed beta-propeller." Structure (London, England : 1993) **3**: 541-549.

Li, S. W., A. L. Sieron, A. Fertala, Y. Hojima, W. V. Arnold and D. J. Prockop (1996). "The C-proteinase that processes procollagens to fibrillar collagens is identical to the protein previously identified as bone morphogenic protein-1." Proc Natl Acad Sci U S A **93**(10): 5127-5130.

Liang, J., E. Liu, Y. Yu, S. Kitajima, T. Koike, Y. Jin, M. Morimoto, K. Hatakeyama, Y. Asada, T. Watanabe, Y. Sasaguri, S. Watanabe and J. Fan (2006). "Macrophage metalloelastase accelerates the progression of atherosclerosis in transgenic rabbits." Circulation **113**(16): 1993-2001.

Lijnen, H. R., J. Silence, B. Van Hoef and D. Collen (1998). "Stromelysin-1 (MMP-3)-independent gelatinase expression and activation in mice." Blood **91**(6): 2045-2053.

Lipfert, J. and S. Doniach (2007). "Small-angle X-ray scattering from RNA, proteins, and protein complexes." Annu Rev Biophys Biomol Struct **36**: 307-327.

Liu, Z., N. Li, L. A. Diaz, M. Shipley, R. M. Senior and Z. Werb (2005). "Synergy between a plasminogen cascade and MMP-9 in autoimmune disease." J Clin Invest **115**(4): 879-887.

Llinas, P., M. H. Le Du, H. Gardsvoll, K. Dano, M. Ploug, B. Gilquin, E. A. Stura and A. Menez (2005). "Crystal structure of the human urokinase plasminogen activator receptor bound to an antagonist peptide." EMBO J **24**(9): 1655-1663.

Lohi, J., C. L. Wilson, J. D. Roby and W. C. Parks (2001). "Epilysin, a novel human matrix metalloproteinase (MMP-28) expressed in testis and keratinocytes and in response to injury." J Biol Chem **276**(13): 10134-10144.

Lovejoy, B., A. R. Welch, S. Carr, C. Luong, C. Broka, R. T. Hendricks, J. A. Campbell, K. A. Walker, R. Martin, H. Van Wart and M. F. Browner (1999). "Crystal structures of MMP-1 and -13 reveal the structural basis for selectivity of collagenase inhibitors." Nat Struct Biol **6**(3): 217-221.

Lu, Y. and L. M. Wahl (2005). "Oxidative stress augments the production of matrix metalloproteinase-1, cyclooxygenase-2, and prostaglandin E2 through enhancement of NF-kappa B activity in lipopolysaccharide-activated human primary monocytes." J Immunol **175**(8): 5423-5429.

Lukashev, M. E. and Z. Werb (1998). "ECM signalling: orchestrating cell behaviour and misbehaviour." Trends Cell Biol **8**(11): 437-441.

Lund, L. R., K. A. Green, A. A. Stoop, M. Ploug, K. Almholt, J. Lilla, B. S. Nielsen, I. J. Christensen, C. S. Craik, Z. Werb, K. Dano and J. Romer (2006). "Plasminogen activation independent of uPA and tPA maintains wound healing in gene-deficient mice." EMBO J **25**(12): 2686-2697.

Lund, L. R., J. Romer, T. H. Bugge, B. S. Nielsen, T. L. Frandsen, J. L. Degen, R. W. Stephens and K. Dano (1999). "Functional overlap between two classes of matrix-degrading proteases in wound healing." EMBO J **18**(17): 4645-4656.

Mackay, A. R., M. Ballin, M. D. Pelina, A. R. Farina, A. M. Nason, J. L. Hartzler and U. P. Thorgeirsson (1992). "Effect of phorbol ester and cytokines on matrix metalloproteinase and tissue inhibitor of metalloproteinase expression in tumor and normal cell lines." Invasion Metastasis **12**(3-4): 168-184.

Makela, M., T. Salo and H. Larjava (1998). "MMP-9 from TNF alpha-stimulated keratinocytes binds to cell membranes and type I collagen: a cause for extended matrix degradation in inflammation?" Biochem Biophys Res Commun **253**(2): 325-335.

Manicone, A. M., T. P. Birkland, M. Lin, T. Betsuyaku, N. van Rooijen, J. Lohi, J. Keski-Oja, Y. Wang, S. J. Skerrett and W. C. Parks (2009). "Epilysin (MMP-28) restrains early macrophage recruitment in Pseudomonas aeruginosa pneumonia." J Immunol **182**(6): 3866-3876.

Manka, S. W., F. Carafoli, R. Visse, D. Bihan, N. Raynal, R. W. Farndale, G. Murphy, J. J. Enghild, E. Hohenester and H. Nagase (2012). "Structural insights into triple-helical collagen cleavage by matrix metalloproteinase 1." Proc Natl Acad Sci U S A **109**(31): 12461-12466.

Mantuano, E., Inoue, G., Li, X., Takahashi, K., Gaultier, A., Gonias, S. L., & Campana, W. M. (2008). The hemopexin domain of matrix metalloproteinase-9 activates cell signaling and promotes migration of schwann cells by binding to low-density lipoprotein receptor-related protein. J Neurosci **28**(45): 11571-11582.

Marchenko, G. N., N. D. Marchenko and A. Y. Strongin (2003). "The structure and regulation of the human and mouse matrix metalloproteinase-21 gene and protein." Biochem J **372**(Pt 2): 503-515.

Marchenko, G. N. and A. Y. Strongin (2001). "MMP-28, a new human matrix metalloproteinase with an unusual cysteine-switch sequence is widely expressed in tumors." Gene **265**(1-2): 87-93.

Marchenko, N. D., G. N. Marchenko, R. N. Weinreb, J. D. Lindsey, A. Kyshtoobayeva, H. C. Crawford and A. Y. Strongin (2004). "Beta-catenin regulates the gene of MMP-26, a novel metalloproteinase expressed both in carcinomas and normal epithelial cells." Int J Biochem Cell Biol **36**(5): 942-956.

Margreiter, G., P. Messner, K. D. Caldwell and K. Bayer (2008). "Size characterization of inclusion bodies by sedimentation field-flow fractionation." J Biotechnol **138**(3-4): 67-73.

Marti, H. P., L. Lee, M. Kashgarian and D. H. Lovett (1994). "Transforming growth factor-beta 1 stimulates glomerular mesangial cell synthesis of the 72-kd type IV collagenase." Am J Pathol **144**(1): 82-94.

Massova, I., L. P. Kotra, R. Fridman and S. Mobashery (1998). "Matrix metalloproteinases: structures, evolution, and diversification." FASEB J **12**: 1075-1095.

Mattu, T. S., L. Royle, J. Langridge, M. R. Wormald, P. E. Van den Steen, J. Van Damme, G. Opdenakker, D. J. Harvey, R. A. Dwek and P. M. Rudd (2000). "O-glycan analysis of natural human neutrophil gelatinase B using a combination of normal phase-HPLC and online tandem mass spectrometry: implications for the domain organization of the enzyme." Biochemistry **39**(51): 15695-15704.

Mazzieri, R., L. Masiero, L. Zanetta, S. Monea, M. Onisto, S. Garbisa and P. Mignatti (1997). "Control of type IV collagenase activity by components of the urokinase-plasmin system: a regulatory mechanism with cell-bound reactants." EMBO J **16**(9): 2319-2332.

McQuibban, G. A., J. H. Gong, J. P. Wong, J. L. Wallace, I. Clark-Lewis and C. M. Overall (2002). "Matrix metalloproteinase processing of monocyte chemoattractant proteins generates CC chemokine receptor antagonists with anti-inflammatory properties in vivo." Blood **100**(4): 1160-1167.

Mehraban, F., S. Y. Kuo, H. Riera, C. Chang and R. W. Moskowitz (1994). "Prostromelysin and procollagenase genes are differentially up-regulated in chondrocytes from the knees of rabbits with experimental osteoarthritis." Arthritis Rheum **37**(8): 1189-1197.

Meli, D. N., S. Christen and S. L. Leib (2003). "Matrix metalloproteinase-9 in pneumococcal meningitis: activation via an oxidative pathway." J Infect Dis **187**(9): 1411-1415.

Menon, B., M. Singh, R. S. Ross, J. N. Johnson and K. Singh (2006). "beta-Adrenergic receptor-stimulated apoptosis in adult cardiac myocytes involves MMP-2-mediated disruption of beta1 integrin signaling and mitochondrial pathway." Am J Physiol Cell Physiol **290**(1): C254-261.

Merck Millipore. (n.d.). "Novagen pET Vector Table - Merck4Biosciences | Merck Millipore United Kingdom." Retrieved 11 Oct, 2012, from http://www.merckmillipore.co.uk/life-science-research/vector-table-novagen-pet-vector-table/c_HdSb.s1O77QAAAEhPqsLdcab.

Mertens, H. D. and D. I. Svergun (2010). "Structural characterization of proteins and complexes using small-angle X-ray solution scattering." J Struct Biol **172**(1): 128-141.

Mitchell, P. G., H. A. Magna, L. M. Reeves, L. L. Lopresti-Morrow, S. A. Yocum, P. J. Rosner, K. F. Geoghegan and J. E. Hambor (1996). "Cloning, expression, and type II collagenolytic activity of matrix metalloproteinase-13 from human osteoarthritic cartilage." J Clin Invest **97**(3): 761-768.

Miyazaki, K., Y. Hattori, F. Umenishi, H. Yasumitsu and M. Umeda (1990). "Purification and characterization of extracellular matrix-degrading metalloproteinase, matrin (pump-1), secreted from human rectal carcinoma cell line." Cancer Res **50**(24): 7758-7764.

Moestrup, S. K., T. L. Holtet, M. Etzerodt, H. C. Thogersen, A. Nykjaer, P. A. Andreasen, H. H. Rasmussen, L. Sottrup-Jensen and J. Gliemann (1993). "Alpha 2-macroglobulin-proteinase complexes, plasminogen activator inhibitor type-1-plasminogen activator complexes, and receptor-associated protein bind to a region of the alpha 2-macroglobulin receptor containing a cluster of eight complement-type repeats." J Biol Chem **268**(18): 13691-13696.

Moller, L. B., J. Pollanen, E. Ronne, N. Pedersen and F. Blasi (1993). "N-linked glycosylation of the ligand-binding domain of the human urokinase receptor contributes to the affinity for its ligand." J Biol Chem **268**(15): 11152-11159.

Molloy, S. S., E. D. Anderson, F. Jean and G. Thomas (1999). "Bi-cycling the furin pathway: from TGN localization to pathogen activation and embryogenesis." Trends Cell Biol **9**(1): 28-35.

Molloy, S. S., P. A. Bresnahan, S. H. Leppla, K. R. Klimpel and G. Thomas (1992). "Human furin is a calcium-dependent serine endoprotease that recognizes the sequence Arg-X-X-Arg and efficiently cleaves anthrax toxin protective antigen." J Biol Chem **267**(23): 16396-16402.

Monea, S., K. Lehti, J. Keski-Oja and P. Mignatti (2002). "Plasmin activates pro-matrix metalloproteinase-2 with a membrane-type 1 matrix metalloproteinase-dependent mechanism." J Cell Physiol **192**(2): 160-170.

Morgunova, E., A. Tuuttila, U. Bergmann, M. Isupov, Y. Lindqvist, G. Schneider and K. Tryggvason (1999). "Structure of human pro-matrix metalloproteinase-2: activation mechanism revealed." Science **284**(5420): 1667-1670.

Morgunova, E., A. Tuuttila, U. Bergmann and K. Tryggvason (2002). "Structural insight into the complex formation of latent matrix metalloproteinase 2 with tissue inhibitor of metalloproteinase 2." Proc Natl Acad Sci U S A **99**(11): 7414-7419.

Mori, H., T. Tomari, N. Koshikawa, M. Kajita, Y. Itoh, H. Sato, H. Tojo, I. Yana and M. Seiki (2002). "CD44 directs membrane-type 1 matrix metalloproteinase to lamellipodia by associating with its hemopexin-like domain." EMBO J **21**(15): 3949-3959.

Morrison, C. J., G. S. Butler, H. F. Bigg, C. R. Roberts, P. D. Soloway and C. M. Overall (2001). "Cellular activation of MMP-2 (gelatinase A) by MT2-MMP occurs via a TIMP-2-independent pathway." J Biol Chem **276**(50): 47402-47410.

Morrison, C. J., G. S. Butler, D. Rodríguez and C. M. Overall (2009). "Matrix metalloproteinase proteomics: substrates, targets, and therapy." Current opinion in cell biology **21**: 645-653.

Moy, F. J., P. K. Chanda, J. M. Chen, S. Cosmi, W. Edris, J. S. Skotnicki, J. Wilhelm and R. Powers (1999). "NMR solution structure of the catalytic fragment of human fibroblast collagenase complexed with a sulfonamide derivative of a hydroxamic acid compound." Biochemistry **38**(22): 7085-7096.

Mucha, A., P. Cuniasse, R. Kannan, F. Beau, A. Yiotakis, P. Basset and V. Dive (1998). "Membrane type-1 matrix metalloprotease and stromelysin-3 cleave more efficiently synthetic substrates containing unusual amino acids in their P1' positions." J Biol Chem **273**(5): 2763-2768.

Muir, D. and M. Manthorpe (1992). "Stromelysin generates a fibronectin fragment that inhibits Schwann cell proliferation." J Cell Biol **116**(1): 177-185.

Muller, D., B. Quantin, M. C. Gesnel, R. Millon-Collard, J. Abecassis and R. Breathnach (1988). "The collagenase gene family in humans consists of at least four members." Biochem J **253**(1): 187-192.

Murphy, G., J. A. Allan, F. Willenbrock, M. I. Cockett, J. P. O'Connell and A. J. Docherty (1992). "The role of the C-terminal domain in collagenase and stromelysin specificity." J Biol Chem **267**(14): 9612-9618.

Murphy, G., M. I. Cockett, P. E. Stephens, B. J. Smith and A. J. Docherty (1987). "Stromelysin is an activator of procollagenase. A study with natural and recombinant enzymes." Biochem J **248**(1): 265-268.

Murphy, G., M. I. Cockett, R. V. Ward and A. J. Docherty (1991). "Matrix metalloproteinase degradation of elastin, type IV collagen and proteoglycan. A quantitative comparison of the activities of 95 kDa and 72 kDa gelatinases, stromelysins-1 and -2 and punctuated metalloproteinase (PUMP)." Biochem J **277 ( Pt 1)**: 277-279.

Murphy, G., Q. Nguyen, M. I. Cockett, S. J. Atkinson, J. A. Allan, C. G. Knight, F. Willenbrock and A. J. Docherty (1994). "Assessment of the role of the fibronectin-like domain of gelatinase A by analysis of a deletion mutant." J Biol Chem **269**(9): 6632-6636.

Murphy, G., J. P. Segain, M. O'Shea, M. Cockett, C. Ioannou, O. Lefebvre, P. Chambon and P. Basset (1993). "The 28-kDa N-terminal domain of mouse stromelysin-3 has the general properties of a weak metalloproteinase." J Biol Chem **268**(21): 15435-15441.

Murphy, G., R. Ward, R. M. Hembry, J. J. Reynolds, K. Kuhn and K. Tryggvason (1989). "Characterization of gelatinase from pig polymorphonuclear leucocytes. A metalloproteinase resembling tumour type IV collagenase." Biochem J **258**(2): 463-472.

Myllyharju, J. (2003). "Prolyl 4-hydroxylases, the key enzymes of collagen biosynthesis." Matrix Biol **22**(1): 15-24.

Mylonas, E. and D. I. Svergun (2007). "Accuracy of molecular mass determination of proteins in solution by small-angle X-ray scattering." Journal of Applied Crystallography **40**(s1): s245-s249.

Myszka, D. G., T. A. Morton, M. L. Doyle and I. M. Chaiken (1997). "Kinetic analysis of a protein antigen-antibody interaction limited by mass transport on an optical biosensor." Biophys Chem **64**(1-3): 127-137.

Naba, A., K. R. Clauser, S. Hoersch, H. Liu, S. A. Carr and R. O. Hynes (2012). "The matrisome: in silico definition and in vivo characterization by proteomics of normal and tumor extracellular matrices." Mol Cell Proteomics **11**(4): M111 014647.

Nagai, N., M. Hosokawa, S. Itohara, E. Adachi, T. Matsushita, N. Hosokawa and K. Nagata (2000). "Embryonic lethality of molecular chaperone hsp47 knockout mice is associated with defects in collagen biosynthesis." J Cell Biol **150**(6): 1499-1506.

Nagase, H., J. J. Enghild, K. Suzuki and G. Salvesen (1990). "Stepwise activation mechanisms of the precursor of matrix metalloproteinase 3 (stromelysin) by proteinases and (4-aminophenyl)mercuric acetate." Biochemistry **29**: 5783-5789.

Nagase, H., C. G. Fields and G. B. Fields (1994). "Design and characterization of a fluorogenic substrate selectively hydrolyzed by stromelysin 1 (matrix metalloproteinase-3)." J Biol Chem **269**(33): 20952-20957.

Nagase, H., K. Suzuki, J. J. Enghild and G. Salvesen (1991). "Stepwise activation mechanisms of the precursors of matrix metalloproteinases 1 (tissue collagenase) and 3 (stromelysin)." Biomed Biochim Acta **50**(4-6): 749-754.

Nagase, H., K. Suzuki, T. Morodomi, J. J. Enghild and G. Salvesen (1992). "Activation mechanisms of the precursors of matrix metalloproteinases 1, 2 and 3." Matrix Suppl **1**: 237-244.

Nagase, H., R. Visse and G. Murphy (2006). "Structure and function of matrix metalloproteinases and TIMPs." Cardiovasc Res **69**(3): 562-573.

Nakahara, H., L. Howard, E. W. Thompson, H. Sato, M. Seiki, Y. Yeh and W. T. Chen (1997). "Transmembrane/cytoplasmic domain-mediated membrane type 1-matrix metalloprotease docking to invadopodia is required for cell invasion." Proc Natl Acad Sci U S A **94**(15): 7959-7964.

Nakamura, H., Y. Fujii, E. Ohuchi, E. Yamamoto and Y. Okada (1998). "Activation of the precursor of human stromelysin 2 and its interactions with other matrix metalloproteinases." Eur J Biochem **253**(1): 67-75.

Nakamura, H., N. Suenaga, K. Taniwaki, H. Matsuki, K. Yonezawa, M. Fujii, Y. Okada and M. Seiki (2004). "Constitutive and induced CD44 shedding by ADAM-like proteases and membrane-type 1 matrix metalloproteinase." Cancer Res **64**(3): 876-882.

Nakamura, Y., T. Gojobori and T. Ikemura (2000). "Codon usage tabulated from international DNA sequence databases: status for the year 2000." Nucleic Acids Res **28**(1): 292.

Nakayama, K. (1997). "Furin: a mammalian subtilisin/Kex2p-like endoprotease involved in processing of a wide variety of precursor proteins." Biochem J **327 ( Pt 3)**: 625-635.

Nerenberg, P. S. and C. M. Stultz (2008). "Differential unfolding of alpha1 and alpha2 chains in type I collagen and collagenolysis." Journal of molecular biology **382**: 246-256.

Netzel-Arnett, S., G. B. Fields, H. Birkedal-Hansen and H. E. Van Wart (1991). "Sequence specificities of human fibroblast and neutrophil collagenases." J Biol Chem **266**(11): 6747-6755.

Nicholson, R., G. Murphy and R. Breathnach (1989). "Human and rat malignant-tumor-associated mRNAs encode stromelysin-like metalloproteinases." Biochemistry **28**(12): 5195-5203.

Nie, J. and D. Pei (2003). "Direct activation of pro-matrix metalloproteinase-2 by leukolysin/membrane-type 6 matrix metalloproteinase/matrix metalloproteinase 25 at the asn(109)-Tyr bond." Cancer Res **63**(20): 6758-6762.

Noel, A., M. Santavicca, I. Stoll, C. L'Hoir, A. Staub, G. Murphy, M. C. Rio and P. Basset (1995). "Identification of structural determinants controlling human and mouse stromelysin-3 proteolytic activities." J Biol Chem **270**(39): 22866-22872.

Nuti, E., T. Tuccinardi and A. Rossello (2007). "Matrix metalloproteinase inhibitors: new challenges in the era of post broad-spectrum inhibitors." Curr Pharm Des **13**(20): 2087-2100.

Nyalendo, C., M. Michaud, E. Beaulieu, C. Roghi, G. Murphy, D. Gingras and R. Beliveau (2007). "Src-dependent phosphorylation of membrane type I matrix metalloproteinase on cytoplasmic tyrosine 573: role in endothelial and tumor cell migration." J Biol Chem **282**(21): 15690-15699.

Ogata, Y., J. J. Enghild and H. Nagase (1992). "Matrix metalloproteinase 3 (stromelysin) activates the precursor for the human matrix metalloproteinase 9." J Biol Chem **267**(6): 3581-3584.

Oh, J., R. Takahashi, S. Kondo, A. Mizoguchi, E. Adachi, R. M. Sasahara, S. Nishimura, Y. Imamura, H. Kitayama, D. B. Alexander, C. Ide, T. P. Horan, T. Arakawa, H. Yoshida, S. Nishikawa, Y. Itoh, M. Seiki, S. Itohara, C. Takahashi and M. Noda (2001). "The membrane-anchored MMP inhibitor RECK is a key regulator of extracellular matrix integrity and angiogenesis." Cell **107**(6): 789-800.

Ohuchi, E., K. Imai, Y. Fujii, H. Sato, M. Seiki and Y. Okada (1997). "Membrane type 1 matrix metalloproteinase digests interstitial collagens and other extracellular matrix macromolecules." J Biol Chem **272**(4): 2446-2451.

Okada, Y., E. D. Harris, Jr. and H. Nagase (1988). "The precursor of a metalloendopeptidase from human rheumatoid synovial fibroblasts. Purification and mechanisms of activation by endopeptidases and 4-aminophenylmercuric acetate." Biochem J **254**(3): 731-741.

Okada, Y., H. Konomi, T. Yada, K. Kimata and H. Nagase (1989). "Degradation of type IX collagen by matrix metalloproteinase 3 (stromelysin) from human rheumatoid synovial cells." FEBS Lett **244**(2): 473-476.

Okada, Y., H. Nagase and E. D. Harris, Jr. (1986). "A metalloproteinase from human rheumatoid synovial fibroblasts that digests connective tissue matrix components. Purification and characterization." J Biol Chem **261**(30): 14245-14255.

Okumura, Y., H. Sato, M. Seiki and H. Kido (1997). "Proteolytic activation of the precursor of membrane type 1 matrix metalloproteinase by human plasmin. A possible cell surface activator." FEBS Lett **402**(2-3): 181-184.

Olson, M. W., M. M. Bernardo, M. Pietila, D. C. Gervasi, M. Toth, L. P. Kotra, I. Massova, S. Mobashery and R. Fridman (2000). "Characterization of the monomeric and dimeric forms of latent and active matrix metalloproteinase-9. Differential rates for activation by stromelysin 1." J Biol Chem **275**(4): 2661-2668.

Orgel, J. P., T. C. Irving, A. Miller and T. J. Wess (2006). "Microfibrillar structure of type I collagen in situ." Proc Natl Acad Sci U S A **103**(24): 9001-9005.

Osenkowski, P., M. Toth and R. Fridman (2004). "Processing, shedding, and endocytosis of membrane type 1-matrix metalloproteinase (MT1-MMP)." J Cell Physiol **200**(1): 2-10.

Ottl, J., D. Gabriel, G. Murphy, V. Knäuper, Y. Tominaga, H. Nagase, M. Kröger, H. Tschesche, W. Bode and L. Moroder (2000). "Recognition and catabolism of synthetic heterotrimeric collagen peptides by matrix metalloproteinases." Chemistry & biology **7**: 119-132.

Overall, C. M. (1994). "Regulation of tissue inhibitor of matrix metalloproteinase expression." Ann N Y Acad Sci **732**: 51-64.

Overall, C. M. (2002). "Molecular Determinants of Metalloproteinase Substrate Specificity." **22**.

Overall, C. M., A. E. King, H. F. Bigg, A. McQuibban, J. Atherstone, D. K. Sam, A. D. Ong, T. T. Lau, U. M. Wallon, Y. A. DeClerck and E. Tam (1999). "Identification of the TIMP-2 binding site on the gelatinase A hemopexin C-domain by site-directed mutagenesis and the yeast two-hybrid system." Ann N Y Acad Sci **878**: 747-753.

Overall, C. M. and C. Lopez-Otin (2002). "Strategies for MMP inhibition in cancer: innovations for the post-trial era." Nat Rev Cancer **2**(9): 657-672.

Overall, C. M., E. Tam, G. A. McQuibban, C. Morrison, U. M. Wallon, H. F. Bigg, A. E. King and C. R. Roberts (2000). "Domain interactions in the gelatinase A.TIMP-2.MT1-MMP activation complex. The ectodomain of the 44-kDa form of membrane type-1 matrix metalloproteinase does not modulate gelatinase A activation." J Biol Chem **275**(50): 39497-39506.

Overall, C. M., J. L. Wrana and J. Sodek (1991). "Transcriptional and post-transcriptional regulation of 72-kDa gelatinase/type IV collagenase by transforming growth factor-beta 1 in human fibroblasts. Comparisons with collagenase and tissue inhibitor of matrix metalloproteinase gene expression." J Biol Chem **266**(21): 14064-14071.

Ozdemir, D., P. S. Hart, O. H. Ryu, S. J. Choi, M. Ozdemir-Karatas, E. Firatli, N. Piesco and T. C. Hart (2005). "MMP20 active-site mutation in hypomaturation amelogenesis imperfecta." J Dent Res **84**(11): 1031-1035.

Page-McCaw, A., A. J. Ewald and Z. Werb (2007). "Matrix metalloproteinases and the regulation of tissue remodelling." Nat Rev Mol Cell Biol **8**(3): 221-233.

Park, A. J., L. M. Matrisian, A. F. Kells, R. Pearson, Z. Y. Yuan and M. Navre (1991). "Mutational analysis of the transin (rat stromelysin) autoinhibitor region demonstrates a role for residues surrounding the "cysteine switch"." J Biol Chem **266**(3): 1584-1590.

Park, H. I., Y. Jin, D. R. Hurst, C. A. Monroe, S. Lee, M. A. Schwartz and Q. X. Sang (2003). "The intermediate S1' pocket of the endometase/matrilysin-2 active site revealed by enzyme inhibition kinetic studies, protein sequence analyses, and homology modeling." J Biol Chem **278**(51): 51646-51653.

Park, H. I., J. Ni, F. E. Gerkema, D. Liu, V. E. Belozerov and Q. X. Sang (2000). "Identification and characterization of human endometase (Matrix metalloproteinase-26) from endometrial tumor." J Biol Chem **275**(27): 20540-20544.

Passoja, K., K. Rautavuoma, L. Ala-Kokko, T. Kosonen and K. I. Kivirikko (1998). "Cloning and characterization of a third human lysyl hydroxylase isoform." Proc Natl Acad Sci U S A **95**(18): 10482-10486.

Patterson, M. L., S. J. Atkinson, V. Knauper and G. Murphy (2001). "Specific collagenolysis by gelatinase A, MMP-2, is determined by the hemopexin domain and not the fibronectin-like domain." FEBS Lett **503**(2-3): 158-162.

Pavlovsky, A. G., M. G. Williams, Q. Z. Ye, D. F. Ortwine, C. F. Purchase, 2nd, A. D. White, V. Dhanaraj, B. D. Roth, L. L. Johnson, D. Hupe, C. Humblet and T. L. Blundell (1999). "X-ray structure of human stromelysin catalytic domain complexed with nonpeptide inhibitors: implications for inhibitor selectivity." Protein Sci **8**(7): 1455-1462.

Pei, D. (1999). "CA-MMP: a matrix metalloproteinase with a novel cysteine array, but without the classic cysteine switch." FEBS Lett **457**(2): 262-270.

Pei, D. (1999). "Identification and characterization of the fifth membrane-type matrix metalloproteinase MT5-MMP." J Biol Chem **274**(13): 8925-8932.

Pei, D., T. Kang and H. Qi (2000). "Cysteine array matrix metalloproteinase (CA-MMP)/MMP-23 is a type II transmembrane matrix metalloproteinase regulated by a single cleavage for both secretion and activation." J Biol Chem **275**(43): 33988-33997.

Pei, D., G. Majmudar and S. J. Weiss (1994). "Hydrolytic inactivation of a breast carcinoma cell-derived serpin by human stromelysin-3." J Biol Chem **269**(41): 25849-25855.

Pei, D. and S. J. Weiss (1995). "Furin-dependent intracellular activation of the human stromelysin-3 zymogen." Nature **375**(6528): 244-247.

Pei, D. and S. J. Weiss (1996). "Transmembrane-deletion mutants of the membrane-type matrix metalloproteinase-1 process progelatinase A and express intrinsic matrix-degrading activity." J Biol Chem **271**(15): 9135-9140.

Pelmenschikov, V. and P. E. Siegbahn (2002). "Catalytic mechanism of matrix metalloproteinases: two-layered ONIOM study." Inorg Chem **41**(22): 5659-5666.

Pendas, A. M., V. Knauper, X. S. Puente, E. Llano, M. G. Mattei, S. Apte, G. Murphy and C. Lopez-Otin (1997). "Identification and characterization of a novel human matrix metalloproteinase with unique structural characteristics, chromosomal location, and tissue distribution." J Biol Chem **272**(7): 4281-4286.

Peppin, G. J. and S. J. Weiss (1986). "Activation of the endogenous metalloproteinase, gelatinase, by triggered human neutrophils." Proc Natl Acad Sci U S A **83**(12): 4322-4326.

Persikov, A. V., J. A. Ramshaw and B. Brodsky (2005). "Prediction of collagen stability from amino acid sequence." J Biol Chem **280**(19): 19343-19349.

Perumal, S., O. Antipova and J. P. R. O. Orgel (2008). "Collagen fibril architecture, domain organization, and triple-helical conformation govern its proteolysis." Proceedings of the National Academy of Sciences of the United States of America **105**: 2824-2829.

Petersen, M. T., Jonson, P. H., & Petersen, S. B. (1999). Amino acid neighbours and detailed conformational analysis of cysteines in proteins. *Protein Eng, 12*(7), 535-548.

Petruska, J. A. and A. J. Hodge (1964). "A Subunit Model for the Tropocollagen Macromolecule." Proc Natl Acad Sci U S A **51**: 871-876.

Pilcher, B. K., J. A. Dumin, B. D. Sudbeck, S. M. Krane, H. G. Welgus and W. C. Parks (1997). "The activity of collagenase-1 is required for keratinocyte migration on a type I collagen matrix." J Cell Biol **137**(6): 1445-1457.

Pinnell, S. R. and G. R. Martin (1968). "The cross-linking of collagen and elastin: enzymatic conversion of lysine in peptide linkage to alpha-aminoadipic-delta-semialdehyde (allysine) by an extract from bone." Proc Natl Acad Sci U S A **61**(2): 708-716.

Plow, E. F., D. E. Freaney, J. Plescia and L. A. Miles (1986). "The plasminogen system and cell surfaces: evidence for plasminogen and urokinase receptors on the same cell type." J Cell Biol **103**(6 Pt 1): 2411-2420.

Puente, X. S., A. M. Pendas, E. Llano, G. Velasco and C. Lopez-Otin (1996). "Molecular cloning of a novel membrane-type matrix metalloproteinase from a human breast carcinoma." Cancer Res **56**(5): 944-949.

Putnam, C. D., M. Hammel, G. L. Hura and J. a. Tainer (2007). "X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution." Quarterly reviews of biophysics **40**: 191-285.

Qiagen (2006). QIAprep Miniprep Handbook**:** 1-51.

Qin, H., J. D. Moellinger, A. Wells, L. J. Windsor, Y. Sun and E. N. Benveniste (1998). "Transcriptional suppression of matrix metalloproteinase-2 gene expression in human astroglioma cells by TNF-alpha and IFN-gamma." J Immunol **161**(12): 6664-6673.

Quantin, B., G. Murphy and R. Breathnach (1989). "Pump-1 cDNA codes for a protein with characteristics similar to those of classical collagenase family members." Biochemistry **28**(13): 5327-5334.

Quesada, V., G. R. Ordonez, L. M. Sanchez, X. S. Puente and C. Lopez-Otin (2009). "The Degradome database: mammalian proteases and diseases of proteolysis." Nucleic Acids Res **37**(Database issue): D239-243.

Rajagopalan, S., X. P. Meng, S. Ramasamy, D. G. Harrison and Z. S. Galis (1996). "Reactive oxygen species produced by macrophage-derived foam cells regulate the activity of vascular matrix metalloproteinases in vitro. Implications for atherosclerotic plaque stability." J Clin Invest **98**(11): 2572-2579.

Ramachandran, G. N. and G. Kartha (1955). "Structure of collagen." Nature **176**(4482): 593-595.

Ramos-DeSimone, N., E. Hahn-Dantona, J. Sipley, H. Nagase, D. L. French and J. P. Quigley (1999). "Activation of matrix metalloproteinase-9 (MMP-9) via a converging plasmin/stromelysin-1 cascade enhances tumor cell invasion." J Biol Chem **274**(19): 13066-13076.

Ramshaw, J. A., N. K. Shah and B. Brodsky (1998). "Gly-X-Y tripeptide frequencies in collagen: a context for host-guest triple-helical peptides." J Struct Biol **122**(1-2): 86-91.

Rawlings, N. D. and A. J. Barrett (1999). "MEROPS: the peptidase database." Nucleic Acids Res **27**(1): 325-331.

Redlitz, A., B. J. Fowler, E. F. Plow and L. A. Miles (1995). "The role of an enolase-related molecule in plasminogen binding to cells." Eur J Biochem **227**(1-2): 407-415.

Reinemer, P., F. Grams, R. Huber, T. Kleine, S. Schnierer, M. Piper, H. Tschesche and W. Bode (1994). "Structural implications for the role of the N terminus in the 'superactivation' of collagenases. A crystallographic study." FEBS Lett **338**(2): 227-233.

Remacle, A. G., D. V. Rozanov, M. Fugere, R. Day and A. Y. Strongin (2006). "Furin regulates the intracellular activation and the uptake rate of cell surface-associated MT1-MMP." Oncogene **25**(41): 5648-5655.

Remacle, A. G., V. S. Golubkov, S. A. Shiryaev, R. Dahl, J. L. Stebbins, A. V. Chernov, A. V. Cheltsov, M. Pellecchia and A. Y. Strongin (2012). "Novel MT1-MMP small-molecule inhibitors based on insights into hemopexin domain function in tumor growth." Cancer Res **72**(9): 2339-2349.

Ricard-Blum, S. (2011). "The collagen family." Cold Spring Harbor perspectives in biology **3**: a004978.

Rich, A. and F. H. Crick (1955). "The structure of collagen." Nature **176**(4489): 915-916.

Rinas, U. and J. E. Bailey (1992). "Protein compositional analysis of inclusion bodies produced in recombinant Escherichia coli." Appl Microbiol Biotechnol **37**(5): 609-614.

Robbins, K. C., L. Summaria, B. Hsieh and R. J. Shah (1967). "The peptide chains of human plasmin. Mechanism of activation of human plasminogen to plasmin." J Biol Chem **242**(10): 2333-2342.

Roeb, E., I. Behrmann, J. Grotzinger, B. Breuer and S. Matern (2000). "An MMP-9 mutant without gelatinolytic activity as a novel TIMP-1-antagonist." FASEB J **14**(12): 1671-1673.

Roldan, A. L., M. V. Cubellis, M. T. Masucci, N. Behrendt, L. R. Lund, K. Dano, E. Appella and F. Blasi (1990). "Cloning and expression of the receptor for human urokinase plasminogen activator, a central molecule in cell surface, plasmin dependent proteolysis." EMBO J **9**(2): 467-474.

Romer, J., T. H. Bugge, C. Pyke, L. R. Lund, M. J. Flick, J. L. Degen and K. Dano (1996). "Impaired wound healing in mice with a disrupted plasminogen gene." Nat Med **2**(3): 287-292.

Roomi, M. W., J. C. Monterrey, T. Kalinovsky, M. Rath and A. Niedzwiecki (2009). "Patterns of MMP-2 and MMP-9 expression in human cancer cell lines." Oncol Rep **21**(5): 1323-1333.

Rosenblum, G., S. R. Cohen, J. Gu, J. Frenkel, R. Sertchook, N. Slack, R. W. Strange, G. Opdenakker and I. Sagi (2007). "Insights into the Structure and Domain Flexibility of Full-Length Pro-Matrix Metalloproteinase-9/Gelatinase B." Structure **1**: 1227-1236.

Rosenblum, G., S. Meroueh, M. Toth, J. F. Fisher, R. Fridman, S. Mobashery and I. Sagi (2007). "Molecular structures and dynamics of the stepwise activation mechanism of a matrix metalloproteinase zymogen: challenging the cysteine switch dogma." Journal of the American Chemical Society **129**: 13566-13574.

Rosenblum, G., P. E. Van den Steen, S. R. Cohen, A. Bitler, D. D. Brand, G. Opdenakker and I. Sagi (2010). "Direct visualization of protease action on collagen triple helical structure." PLoS One **5**(6): e11043.

Rosenfeld, S. A., O. H. Ross, M. C. Hillman, J. I. Corman and R. L. Dowling (1996). "Production and purification of human fibroblast collagenase (MMP-1) expressed in the methylotrophic yeast Pichia pastoris." Protein Expr Purif **7**(4): 423-430.

Rowan, A. D. and D. A. Young (2007). "Collagenase gene regulation by pro-inflammatory cytokines in cartilage." Front Biosci **12**: 536-550.

Rozanov, D. V., B. Ghebrehiwet, B. Ratnikov, E. Z. Monosov, E. I. Deryugina and A. Y. Strongin (2002). "The cytoplasmic tail peptide sequence of membrane type-1 matrix metalloproteinase (MT1-MMP) directly binds to gC1qR, a compartment-specific chaperone-like regulatory protein." FEBS Lett **527**(1-3): 51-57.

Ryu, O. H., A. G. Fincham, C. C. Hu, C. Zhang, Q. Qian, J. D. Bartlett and J. P. Simmer (1999). "Characterization of recombinant pig enamelysin activity and cleavage of recombinant pig and mouse amelogenins." J Dent Res **78**(3): 743-750.

Saarialho-Kere, U. K., A. P. Pentland, H. Birkedal-Hansen, W. C. Parks and H. G. Welgus (1994). "Distinct populations of basal keratinocytes express stromelysin-1 and stromelysin-2 in chronic wounds." J Clin Invest **94**(1): 79-88.

Saarinen, J., H. G. Welgus, C. A. Flizar, N. Kalkkinen and J. Helin (1999). "N-glycan structures of matrix metalloproteinase-1 derived from human fibroblasts and from HT-1080 fibrosarcoma cells." Eur J Biochem **259**(3): 829-840.

Saffarian, S., I. E. Collier, B. L. Marmer, E. L. Elson and G. Goldberg (2004). "Interstitial collagenase is a Brownian ratchet driven by proteolysis of collagen." Science **306**(5693): 108-111.

Salsas-Escat, R., P. S. Nerenberg and C. M. Stultz (2010). "Cleavage site specificity and conformational selection in type I collagen degradation." Biochemistry.

Sanchez-Lopez, R., C. M. Alexander, O. Behrendtsen, R. Breathnach and Z. Werb (1993). "Role of zinc-binding- and hemopexin domain-encoded sequences in the substrate specificity of collagenase and stromelysin-2 as revealed by chimeric proteins." J Biol Chem **268**(10): 7238-7247.

Sanchez-Lopez, R., R. Nicholson, M. C. Gesnel, L. M. Matrisian and R. Breathnach (1988). "Structure-function relationships in the collagenase family member transin." J Biol Chem **263**(24): 11892-11899.

Sang, Q. X., H. Birkedal-Hansen and H. E. Van Wart (1995). "Proteolytic and non-proteolytic activation of human neutrophil progelatinase B." Biochim Biophys Acta **1251**(2): 99-108.

Santala, A., J. Saarinen, P. Kovanen and P. Kuusela (1999). "Activation of interstitial collagenase, MMP-1, by Staphylococcus aureus cells having surface-bound plasmin: a novel role of plasminogen receptors of bacteria." FEBS Lett **461**(3): 153-156.

Santavicca, M., A. Noel, H. Angliker, I. Stoll, J. P. Segain, P. Anglard, M. Chretien, N. Seidah and P. Basset (1996). "Characterization of structural determinants and molecular mechanisms involved in pro-stromelysin-3 activation by 4-aminophenylmercuric acetate and furin-type convertases." Biochem J **315 ( Pt 3)**: 953-958.

Sapna, G., S. Gokul and K. Bagri-Manjrekar (2013). "Matrix metalloproteinases and periodontal diseases." Oral Dis.

Sarkar, S. K., B. Marmer, G. Goldberg and K. C. Neuman (2012). "Single-molecule tracking of collagenase on native type I collagen fibrils reveals degradation mechanism." Curr Biol **22**(12): 1047-1056.

Sasisekharan, R., R. Raman and V. Prabhakar (2006). "Glycomics approach to structure-function relationships of glycosaminoglycans." Annu Rev Biomed Eng **8**: 181-231.

Sato, H., T. Kinoshita, T. Takino, K. Nakayama and M. Seiki (1996). "Activation of a recombinant membrane type 1-matrix metalloproteinase (MT1-MMP) by furin and its interaction with tissue inhibitor of metalloproteinases (TIMP)-2." FEBS Lett **393**(1): 101-104.

Sato, T., M. del Carmen Ovejero, P. Hou, A. M. Heegaard, M. Kumegawa, N. T. Foged and J. M. Delaisse (1997). "Identification of the membrane-type matrix metalloproteinase MT1-MMP in osteoclasts." J Cell Sci **110 ( Pt 5)**: 589-596.

Satoh, M., K. Hirayoshi, S. Yokota, N. Hosokawa and K. Nagata (1996). "Intracellular interaction of collagen-specific stress protein HSP47 with newly synthesized procollagen." J Cell Biol **133**(2): 469-483.

Sauk, J. J., T. Smith, K. Norris and L. Ferreira (1994). "Hsp47 and the translation-translocation machinery cooperate in the production of alpha 1(I) chains of type I procollagen." J Biol Chem **269**(6): 3941-3946.

Saus, J., S. Quinones, Y. Otani, H. Nagase, E. D. Harris, Jr. and M. Kurkinen (1988). "The complete primary structure of human matrix metalloproteinase-3. Identity with stromelysin." J Biol Chem **263**(14): 6742-6745.

Schettler, A., H. Thorn, B. M. Jockusch and H. Tschesche (1991). "Release of proteinases from stimulated polymorphonuclear leukocytes. Evidence for subclasses of the main granule types and their association with cytoskeletal components." Eur J Biochem **197**(1): 197-202.

Schnierer, S., T. Kleine, T. Gote, A. Hillemann, V. Knauper and H. Tschesche (1993). "The recombinant catalytic domain of human neutrophil collagenase lacks type I collagen substrate specificity." Biochem Biophys Res Commun **191**(2): 319-326.

Schrodinger, LLC. (2010). The PyMOL Molecular Graphics System, Version 1.5.0.4.

Seandel, M., K. Noack-Kunnmann, D. Zhu, R. T. Aimes and J. P. Quigley (2001). "Growth factor-induced angiogenesis in vivo requires specific cleavage of fibrillar type I collagen." Blood **97**(8): 2323-2332.

Sedlacek, R., S. Mauch, B. Kolb, C. Schatzlein, H. Eibel, H. H. Peter, J. Schmitt and U. Krawinkel (1998). "Matrix metalloproteinase MMP-19 (RASI-1) is expressed on the surface of activated peripheral blood mononuclear cells and is detected as an autoantigen in rheumatoid arthritis." Immunobiology **198**(4): 408-423.

Seltzer, J. L., S. A. Adams, G. A. Grant and A. Z. Eisen (1981). "Purification and properties of a gelatin-specific neutral protease from human skin." J Biol Chem **256**(9): 4662-4668.

Shah, V., S. Kumar and K. A. Zirvi (1994). "Metastasis of human colon tumor cells in vivo: correlation with the overexpression of plasminogen activators and 72 kDa gelatinase." In Vivo **8**(3): 321-326.

Shapiro, S. D., D. K. Kobayashi and T. J. Ley (1993). "Cloning and characterization of a unique elastolytic metalloproteinase produced by human alveolar macrophages." J Biol Chem **268**(32): 23824-23829.

Sharp, P. M., E. Cowe, D. G. Higgins, D. C. Shields, K. H. Wolfe and F. Wright (1988). "Codon usage patterns in Escherichia coli, Bacillus subtilis, Saccharomyces cerevisiae, Schizosaccharomyces pombe, Drosophila melanogaster and Homo sapiens; a review of the considerable within-species diversity." Nucleic Acids Res **16**(17): 8207-8211.

Shimada, T., H. Nakamura, E. Ohuchi, Y. Fujii, Y. Murakami, H. Sato, M. Seiki and Y. Okada (1999). "Characterization of a truncated recombinant form of human membrane type 3 matrix metalloproteinase." Eur J Biochem **262**(3): 907-914.

Shimaoka, M., Lu, C., Palframan, R. T., von Andrian, U. H., McCormack, A., Takagi, J., & Springer, T. A. (2001). Reversibly locking a protein fold in an active conformation with a disulfide bond: integrin alphaL I domains with high affinity and antagonist activity in vivo. *Proc Natl Acad Sci U S A, 98*(11), 6009-6014. doi: 10.1073/pnas.101130498

Siegel, R. C. (1974). "Biosynthesis of collagen crosslinks: increased activity of purified lysyl oxidase with reconstituted collagen fibrils." Proc Natl Acad Sci U S A **71**(12): 4826-4830.

Sievers, F., A. Wilm, D. Dineen, T. J. Gibson, K. Karplus, W. Li, R. Lopez, H. McWilliam, M. Remmert, J. Soding, J. D. Thompson and D. G. Higgins (2011). "Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega." Mol Syst Biol **7**: 539.

Silletti, S., T. Kessler, J. Goldberg, D. L. Boger and D. A. Cheresh (2001). "Disruption of matrix metalloproteinase 2 binding to integrin alpha vbeta 3 by an organic molecule inhibits angiogenesis and tumor growth in vivo." Proc Natl Acad Sci U S A **98**(1): 119-124.

Sires, U. I., G. L. Griffin, T. J. Broekelmann, R. P. Mecham, G. Murphy, A. E. Chung, H. G. Welgus and R. M. Senior (1993). "Degradation of entactin by matrix metalloproteinases. Susceptibility to matrilysin and identification of cleavage sites." J Biol Chem **268**(3): 2069-2074.

Sober, H. A. (1968). Handbook of Biochemistry: Selected Data for Molecular Biology. Cleveland, OH, Chemical Rubber Co.

Solberg, H., J. Rinkenberger, K. Dano, Z. Werb and L. R. Lund (2003). "A functional overlap of plasminogen and MMPs regulates vascularization during placental development." Development **130**(18): 4439-4450.

Solomon, A., D. Q. Li, S. B. Lee and S. C. Tseng (2000). "Regulation of collagenase, stromelysin, and urokinase-type plasminogen activator in primary pterygium body fibroblasts by inflammatory cytokines." Invest Ophthalmol Vis Sci **41**(8): 2154-2163.

Sorensen, H. P. and K. K. Mortensen (2005). "Advanced genetic strategies for recombinant protein expression in Escherichia coli." J Biotechnol **115**(2): 113-128.

Sottrup-Jensen, L. and H. Birkedal-Hansen (1989). "Human fibroblast collagenase-alpha-macroglobulin interactions. Localization of cleavage sites in the bait regions of five mammalian alpha-macroglobulins." J Biol Chem **264**(1): 393-401.

Sottrup-Jensen, L. and H. Birkedal-Hansen (1992). "Localization of cleavage sites for human fibroblast collagenase in the bait region of five mammalian alpha-macroglobulins." Matrix Suppl **1**: 263-268.

Springman, E. B., E. L. Angleton, H. Birkedal-Hansen and H. E. Van Wart (1990). "Multiple modes of activation of latent human fibroblast collagenase: evidence for the role of a Cys73 active-site zinc complex in latency and a "cysteine switch" mechanism for activation." Proc Natl Acad Sci U S A **87**(1): 364-368.

Stahle-Backdahl, M., B. Sandstedt, K. Bruce, A. Lindahl, M. G. Jimenez, J. A. Vega and C. Lopez-Otin (1997). "Collagenase-3 (MMP-13) is expressed during human fetal ossification and re-expressed in postnatal bone remodeling and in rheumatoid arthritis." Lab Invest **76**(5): 717-728.

Steele, D. L., O. El-Kabbani, P. Dunten, L. J. Windsor, R. U. Kammlott, R. L. Crowther, C. Michoud, J. A. Engler and J. J. Birktoft (2000). "Expression, characterization and structure determination of an active site mutant (Glu202-Gln) of mini-stromelysin-1." Protein Eng **13**(6): 397-405.

Steffensen, B., U. M. Wallon and C. M. Overall (1995). "Extracellular matrix binding properties of recombinant fibronectin type II-like modules of human 72-kDa gelatinase/type IV collagenase. High affinity binding to native type I collagen but not native type IV collagen." J Biol Chem **270**(19): 11555-11566.

Stricker, T. P., J. A. Dumin, S. K. Dickeson, L. Chung, H. Nagase, W. C. Parks and S. A. Santoro (2001). "Structural analysis of the alpha(2) integrin I domain/procollagenase-1 (matrix metalloproteinase-1) interaction." J Biol Chem **276**(31): 29375-29381.

Stricklin, G. P., J. J. Jeffrey, W. T. Roswit and a. Z. Eisen (1983). "Human skin fibroblast procollagenase: mechanisms of activation by organomercurials and trypsin." Biochemistry **22**: 61-68.

Strongin, A. Y., I. Collier, G. Bannikov, B. L. Marmer, G. A. Grant and G. I. Goldberg (1995). "Mechanism of cell surface activation of 72-kDa type IV collagenase. Isolation of the activated form of the membrane metalloprotease." J Biol Chem **270**(10): 5331-5338.

Studier, F. W. and B. A. Moffatt (1986). "Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes." J Mol Biol **189**(1): 113-130.

Stultz, C. M. (2002). "Localized Unfolding of Collagen Explains Collagenase Cleavage Near Imino-poor Sites." Journal of Molecular Biology **319**: 997-1003.

Su, G., S. A. Blaine, D. Qiao and A. Friedl (2008). "Membrane type 1 matrix metalloproteinase-mediated stromal syndecan-1 shedding stimulates breast carcinoma cell proliferation." Cancer Res **68**(22): 9558-9565.

Suzuki, A. S., Y. Tadano, T. Yamamoto, S. I. Abe and T. Tajima (2001). "Expression of a novel matrix metalloproteinase gene during Cynops early embryogenesis." Biochem Biophys Res Commun **288**(2): 380-384.

Suzuki, K., J. J. Enghild, T. Morodomi, G. Salvesen and H. Nagase (1990). "Mechanisms of activation of tissue procollagenase by matrix metalloproteinase 3 (stromelysin)." Biochemistry **29**(44): 10261-10270.

Suzuki, K., J. J. Enghild, T. Morodomi, G. Salvesen and H. Nagase (1990). "Mechanisms of activation of tissue procollagenase by matrix metalloproteinase 3 (stromelysin)." Biochemistry **29**: 10261-10270.

Suzuki, K., M. Lees, G. F. Newlands, H. Nagase and D. E. Woolley (1995). "Activation of precursors for matrix metalloproteinases 1 (interstitial collagenase) and 3 (stromelysin) by rat mast-cell proteinases I and II." Biochem J **305 ( Pt 1)**: 301-306.

Svergun, D. (1992). "Determination of the regularization parameter in indirect-transform methods using perceptual criteria." Journal of Applied Crystallography **25**(4): 495-503.

Svergun, D., C. Barberato and M. H. J. Koch (1995). "CRYSOL - a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates." Journal of Applied Crystallography **28**(6): 768-773.

Svergun, D. I. and M. H. Koch (2002). "Advances in structure analysis using small-angle scattering in solution." Curr Opin Struct Biol **12**(5): 654-660.

Takahashi, C., Z. Sheng, T. P. Horan, H. Kitayama, M. Maki, K. Hitomi, Y. Kitaura, S. Takai, R. M. Sasahara, A. Horimoto, Y. Ikawa, B. J. Ratzkin, T. Arakawa and M. Noda (1998). "Regulation of matrix metalloproteinase-9 and inhibition of tumor invasion by the membrane-anchored glycoprotein RECK." Proc Natl Acad Sci U S A **95**(22): 13221-13226.

Takino, T., H. Sato, A. Shinagawa and M. Seiki (1995). "Identification of the second membrane-type matrix metalloproteinase (MT-MMP-2) gene from a human placenta cDNA library. MT-MMPs form a unique membrane-type subclass in the MMP family." J Biol Chem **270**(39): 23013-23020.

Tallant, C., A. Marrero and F. X. Gomis-Ruth (2010). "Matrix metalloproteinases: fold and function of their catalytic domains." Biochim Biophys Acta **1803**(1): 20-28.

Tam, E. M., Y. I. Wu, G. S. Butler, M. S. Stack and C. M. Overall (2002). "Collagen binding properties of the membrane type-1 matrix metalloproteinase (MT1-MMP) hemopexin C domain. The ectodomain of the 44-kDa autocatalytic product of MT1-MMP inhibits cell invasion by disrupting native type I collagen cleavage." J Biol Chem **277**(41): 39005-39014.

Tam, E. M., T. R. Moore, G. S. Butler and C. M. Overall (2004). "Characterization of the distinct collagen binding, helicase and cleavage mechanisms of matrix metalloproteinase 2 and 14 (gelatinase A and MT1-MMP): the differential roles of the MMP hemopexin c domains and the MMP-2 fibronectin type II modules in collagen triple helicase activities." J Biol Chem **279**(41): 43336-43344.

Templeton, N. S., P. D. Brown, A. T. Levy, I. M. Margulies, L. A. Liotta and W. G. Stetler-Stevenson (1990). "Cloning and characterization of human tumor cell interstitial collagenase." Cancer Res **50**(17): 5431-5437.

Tetlow, L. C., D. J. Adlam and D. E. Woolley (2001). "Matrix metalloproteinase and proinflammatory cytokine production by chondrocytes of human osteoarthritic cartilage: associations with degenerative changes." Arthritis Rheum **44**(3): 585-594.

Thermo Scientific. (2009). "Carbodiimide Crosslinker Chemistry." Retrieved 23 July 2013, from: http://www.piercenet.com/browse.cfm?fldID=F3305493-0FBC-93DA-2720-4412D198A9C9.

Thomas, G. (2002). "Furin at the cutting edge: from protein traffic to embryogenesis and disease." Nat Rev Mol Cell Biol **3**(10): 753-766.

Tochowicz, A., P. Goettig, R. Evans, R. Visse, Y. Shitomi, R. Palmisano, N. Ito, K. Richter, K. Maskos, D. Franke, D. Svergun, H. Nagase, W. Bode and Y. Itoh (2010). "The dimer interface of the MT1-MMP hemopexin domain: crystal structure and biological functions." The Journal of biological chemistry.

Tochowicz, A., P. Goettig, R. Evans, R. Visse, Y. Shitomi, R. Palmisano, N. Ito, K. Richter, K. Maskos, D. Franke, D. Svergun, H. Nagase, W. Bode and Y. Itoh (2011). "The dimer interface of the membrane type 1 matrix metalloproteinase hemopexin domain: crystal structure and biological functions." J Biol Chem **286**(9): 7587-7600.

Troeberg, L., M. Tanaka, R. Wait, Y. E. Shi, K. Brew and H. Nagase (2002). "E. coli expression of TIMP-4 and comparative kinetic studies with TIMP-1 and TIMP-2: insights into the interactions of TIMPs and matrix metalloproteinase 2 (gelatinase A)." Biochemistry **41**(50): 15025-15035.

Tsukada, H. and T. Pourmotabbed (2002). "Unexpected crucial role of residue 272 in substrate specificity of fibroblast collagenase." J Biol Chem **277**(30): 27378-27384.

Tuinstra, R. L., Peterson, F. C., Elgin, E. S., Pelzek, A. J., & Volkman, B. F. (2007). An engineered second disulfide bond restricts lymphotactin/XCL1 to a chemokine-like conformation with XCR1 agonist activity. *Biochemistry, 46*(10), 2564-2573. doi: 10.1021/bi602365d

Turk, B. E., D. H. Lee, Y. Yamakoshi, A. Klingenhoff, E. Reichenberger, J. T. Wright, J. P. Simmer, J. A. Komisarof, L. C. Cantley and J. D. Bartlett (2006). "MMP-20 is predominately a tooth-specific enzyme with a deep catalytic pocket that hydrolyzes type V collagen." Biochemistry **45**(12): 3863-3874.

Uria, J. A. and C. Lopez-Otin (2000). "Matrilysin-2, a new matrix metalloproteinase expressed in human tumors and showing the minimal domain organization required for secretion, latency, and activity." Cancer Res **60**(17): 4745-4751.

Valentin, F., J. L. Bueb, P. Kieffer, E. Tschirhart and J. Atkinson (2005). "Oxidative stress activates MMP-2 in cultured human coronary smooth muscle cells." Fundam Clin Pharmacol **19**(6): 661-667.

Valtavaara, M., C. Szpirer, J. Szpirer and R. Myllyla (1998). "Primary structure, tissue distribution, and chromosomal localization of a novel isoform of lysyl hydroxylase (lysyl hydroxylase 3)." J Biol Chem **273**(21): 12881-12886.

Van den Steen, P. E., P. Proost, A. Wuyts, J. Van Damme and G. Opdenakker (2000). "Neutrophil gelatinase B potentiates interleukin-8 tenfold by aminoterminal processing, whereas it degrades CTAP-III, PF-4, and GRO-alpha and leaves RANTES and MCP-2 intact." Blood **96**(8): 2673-2681.

Van Wart, H. E. and H. Birkedal-Hansen (1990). "The cysteine switch: a principle of regulation of metalloproteinase activity with potential applicability to the entire matrix metalloproteinase gene family." Proc Natl Acad Sci U S A **87**(14): 5578-5582.

Velasco, G., A. M. Pendas, A. Fueyo, V. Knauper, G. Murphy and C. Lopez-Otin (1999). "Cloning and characterization of human MMP-23, a new matrix metalloproteinase predominantly expressed in reproductive tissues and lacking conserved domains in other family members." J Biol Chem **274**(8): 4570-4576.

Verzijl, N., J. DeGroot, S. R. Thorpe, R. A. Bank, J. N. Shaw, T. J. Lyons, J. W. Bijlsma, F. P. Lafeber, J. W. Baynes and J. M. TeKoppele (2000). "Effect of collagen turnover on the accumulation of advanced glycation end products." J Biol Chem **275**(50): 39027-39031.

Visse, R. and H. Nagase (2003). "Matrix metalloproteinases and tissue inhibitors of metalloproteinases: structure, function, and biochemistry." Circ Res **92**(8): 827-839.

Volkov, V. V. and D. I. Svergun (2003). "Uniqueness of ab initio shape determination in small-angle scattering." Journal of Applied Crystallography **36**(3 Part 1): 860-864.

Vranken, W. F., W. Boucher, T. J. Stevens, R. H. Fogh, A. Pajon, M. Llinas, E. L. Ulrich, J. L. Markley, J. Ionides and E. D. Laue (2005). "The CCPN data model for NMR spectroscopy: development of a software pipeline." Proteins **59**(4): 687-696.

Walker, J. A., S. S. Molloy, G. Thomas, T. Sakaguchi, T. Yoshida, T. M. Chambers and Y. Kawaoka (1994). "Sequence specificity of furin, a proprotein-processing endoprotease, for the hemagglutinin of a virulent avian influenza virus." J Virol **68**(2): 1213-1218.

Wallon, U. M. and C. M. Overall (1997). "The hemopexin-like domain (C domain) of human gelatinase A (matrix metalloproteinase-2) requires Ca2+ for fibronectin and heparin binding. Binding properties of recombinant gelatinase A C domain to extracellular matrix and basement membrane components." J Biol Chem **272**(11): 7473-7481.

Walsh, G. (2006). "Biopharmaceutical benchmarks 2006." Nat Biotechnol **24**(7): 769-776.

Wang, W. M., S. Lee, B. M. Steiglitz, I. C. Scott, C. C. Lebares, M. L. Allen, M. C. Brenner, K. Takahara and D. S. Greenspan (2003). "Transforming growth factor-beta induces secretion of activated ADAMTS-2. A procollagen III N-proteinase." J Biol Chem **278**(21): 19549-19557.

Waterhouse, A. M., J. B. Procter, D. M. Martin, M. Clamp and G. J. Barton (2009). "Jalview Version 2--a multiple sequence alignment editor and analysis workbench." Bioinformatics **25**(9): 1189-1191.

Welgus, H. G., J. J. Jeffrey, G. P. Stricklin, W. T. Roswit and A. Z. Eisen (1980). "Characteristics of the action of human skin fibroblast collagenase on fibrillar collagen." J Biol Chem **255**(14): 6806-6813.

Welgus, H. G., J. J. Jeffrey and A. Z. Eisen (1981). "The collagen substrate specificity of human skin fibroblast collagenase." J Biol Chem **256**(18): 9511-9515.

Welgus, H. G., J. J. Jeffrey, A. Z. Eisen, W. T. Roswit and G. P. Stricklin (1985). "Human skin fibroblast collagenase: interaction with substrate and inhibitor." Coll Relat Res **5**(2): 167-179.

Wewer, U. M., M. Morgelin, P. Holck, J. Jacobsen, M. C. Lydolph, A. H. Johnsen, M. Kveiborg and R. Albrechtsen (2006). "ADAM12 is a four-leafed clover: the excised prodomain remains bound to the mature enzyme." J Biol Chem **281**(14): 9418-9422.

Whitham, S. E., G. Murphy, P. Angel, H. J. Rahmsdorf, B. J. Smith, A. Lyons, T. J. Harris, J. J. Reynolds, P. Herrlich and A. J. Docherty (1986). "Comparison of human stromelysin and collagenase by cloning and sequence analysis." Biochem J **240**(3): 913-916.

Wilhelm, S. M., I. E. Collier, A. Kronberger, A. Z. Eisen, B. L. Marmer, G. A. Grant, E. A. Bauer and G. I. Goldberg (1987). "Human skin fibroblast stromelysin: structure, glycosylation, substrate specificity, and differential expression in normal and tumorigenic cells." Proc Natl Acad Sci U S A **84**(19): 6725-6729.

Wilhelm, S. M., I. E. Collier, B. L. Marmer, A. Z. Eisen, G. A. Grant and G. I. Goldberg (1989). "SV40-transformed human lung fibroblasts secrete a 92-kDa type IV collagenase which is identical to that secreted by normal human macrophages." J Biol Chem **264**(29): 17213-17221.

Wilhelm, S. M., A. Z. Eisen, M. Teter, S. D. Clark, A. Kronberger and G. Goldberg (1986). "Human fibroblast collagenase: glycosylation and tissue-specific levels of enzyme synthesis." Proc Natl Acad Sci U S A **83**(11): 3756-3760.

Will, H., S. J. Atkinson, G. S. Butler, B. Smith and G. Murphy (1996). "The soluble catalytic domain of membrane type 1 matrix metalloproteinase cleaves the propeptide of progelatinase A and initiates autoproteolytic activation. Regulation by TIMP-2 and TIMP-3." J Biol Chem **271**(29): 17119-17123.

Will, H. and B. Hinzmann (1995). "cDNA sequence and mRNA tissue distribution of a novel human matrix metalloproteinase with a potential transmembrane segment." Eur J Biochem **231**(3): 602-608.

Williams, A., S. V. Hill and I. T. Ibrahim (1981). "Improved spectrophotometric methods for the assay of carbodiimides." Anal Biochem **114**(1): 173-176.

Williams, H., J. L. Johnson, C. L. Jackson, S. J. White and S. J. George (2010). "MMP-7 mediates cleavage of N-cadherin and promotes smooth muscle cell apoptosis." Cardiovasc Res **87**(1): 137-146.

Williamson, R. A., F. A. Marston, S. Angal, P. Koklitis, M. Panico, H. R. Morris, A. F. Carne, B. J. Smith, T. J. Harris and R. B. Freedman (1990). "Disulphide bond assignment in human tissue inhibitor of metalloproteinases (TIMP)." Biochem J **268**(2): 267-274.

Williamson, R. A., Muskett, F. W., Howard, M. J., Freedman, R. B., & Carr, M. D. (1999). The effect of matrix metalloproteinase complex formation on the conformational mobility of tissue inhibitor of metalloproteinases-2 (TIMP-2). J Biol Chem **274**(52), 37226-37232

Windsor, L. J., H. Birkedal-Hansen, B. Birkedal-Hansen and J. A. Engler (1991). "An internal cysteine plays a role in the maintenance of the latency of human fibroblast collagenase." Biochemistry **30**(3): 641-647.

Windsor, L. J., M. K. Bodden, B. Birkedal-Hansen, J. A. Engler and H. Birkedal-Hansen (1994). "Mutational analysis of residues in and around the active site of human fibroblast-type collagenase." J Biol Chem **269**(42): 26201-26207.

Windsor, L. J., H. Grenett, B. Birkedal-Hansen, M. K. Bodden, J. A. Engler and H. Birkedal-Hansen (1993). "Cell type-specific regulation of SL-1 and SL-2 genes. Induction of the SL-2 gene but not

the SL-1 gene by human keratinocytes in response to cytokines and phorbolesters." J Biol Chem **268**(23): 17341-17347.

Woessner, J. F., Jr. and C. J. Taplin (1988). "Purification and properties of a small latent matrix metalloproteinase of the rat uterus." J Biol Chem **263**(32): 16918-16925.

Wu, B., S. P. Crampton and C. C. Hughes (2007). "Wnt signaling induces matrix metalloproteinase expression and regulates T cell transmigration." Immunity **26**(2): 227-239.

Xiao, J., R. M. Addabbo, J. L. Lauer, G. B. Fields and J. Baum (2010). "Local conformation and dynamics of isoleucine in the collagenase cleavage site provides recognition signal for matrix metalloproteinases." The Journal of biological chemistry: 1-22.

Xu, X., Y. Wang, J. L. Lauer-Fields, G. B. Fields and B. Steffensen (2004). "Contributions of the MMP-2 collagen binding domain to gelatin cleavage. Substrate binding via the collagen binding domain is required for hydrolysis of gelatin but not short peptides." Matrix Biol **23**(3): 171-181.

Yamada, S., K. Y. Wang, A. Tanimoto, J. Fan, S. Shimajiri, S. Kitajima, M. Morimoto, M. Tsutsui, T. Watanabe, K. Yasumoto and Y. Sasaguri (2008). "Matrix metalloproteinase 12 accelerates the initiation of atherosclerosis and stimulates the progression of fatty streaks to fibrous plaques in transgenic rabbits." Am J Pathol **172**(5): 1419-1429.

Yamamoto, M., H. Tsujishita, N. Hori, Y. Ohishi, S. Inoue, S. Ikeda and Y. Okada (1998). "Inhibition of membrane-type 1 matrix metalloproteinase by hydroxamate inhibitors: an examination of the subsite pocket." J Med Chem **41**(8): 1209-1217.

Yan, C. and D. D. Boyd (2006). "Regulation of Matrix Metalloproteinase Gene Expression." Journal of Cellular Physiology: 19-26.

Yana, I. and S. J. Weiss (2000). "Regulation of membrane type-1 matrix metalloproteinase activation by proprotein convertases." Mol Biol Cell **11**(7): 2387-2401.

Yang, M., M. T. Murray and M. Kurkinen (1997). "A novel matrix metalloproteinase gene (XMMP) encoding vitronectin-like motifs is transiently expressed in Xenopus laevis early embryo development." J Biol Chem **272**(21): 13527-13533.

Yoshihara, Y., H. Nakamura, K. Obata, H. Yamada, T. Hayakawa, K. Fujikawa and Y. Okada (2000). "Matrix metalloproteinases and tissue inhibitors of metalloproteinases in synovial fluids from patients with rheumatoid arthritis or osteoarthritis." Ann Rheum Dis **59**(6): 455-461.

Young, D. A., A. D. Rowan and I. M. Clark (2010). Matrix Metalloproteinase Protocols, Humana Press.

Yu, Q. and I. Stamenkovic (1999). "Localization of matrix metalloproteinase 9 to the cell surface provides a mechanism for CD44-mediated tumor invasion." Genes Dev **13**(1): 35-48.

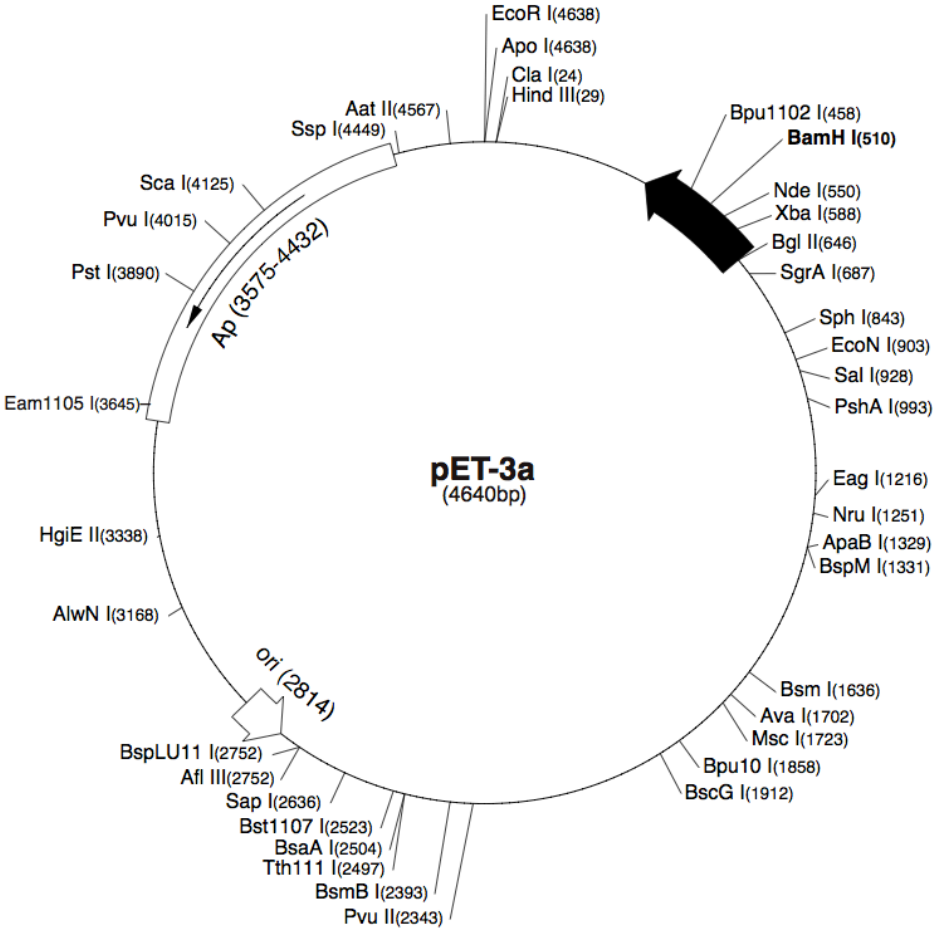
Yu, W. H., J. F. Woessner, Jr., J. D. McNeish and I. Stamenkovic (2002). "CD44 anchors the assembly of matrilysin/MMP-7 with heparin-binding epidermal growth factor precursor and ErbB4 and regulates female reproductive organ remodeling." Genes Dev **16**(3): 307-323.

Zhang, X., Gonnella, N. C., Koehn, J., Pathak, N., Ganu, V., Melton, R., . . . Nam, K. Y. (2000). Solution structure of the catalytic domain of human collagenase-3 (MMP-13) complexed to a potent non-peptidic sulfonamide inhibitor: binding comparison with stromelysin-1 and collagenase-1. *J Mol Biol, 301*(2), 513-524. doi: 10.1006/jmbi.2000.3988

Zhao, H., M. M. Bernardo, P. Osenkowski, A. Sohail, D. Pei, H. Nagase, M. Kashiwagi, P. D. Soloway, Y. A. DeClerck and R. Fridman (2004). "Differential inhibition of membrane type 3 (MT3)-matrix metalloproteinase (MMP) and MT1-MMP by tissue inhibitor of metalloproteinase (TIMP)-2 and TIMP-3 rgulates pro-MMP-2 activation." J Biol Chem **279**(10): 8592-8601.

APPENDIX 1

## pET-3a

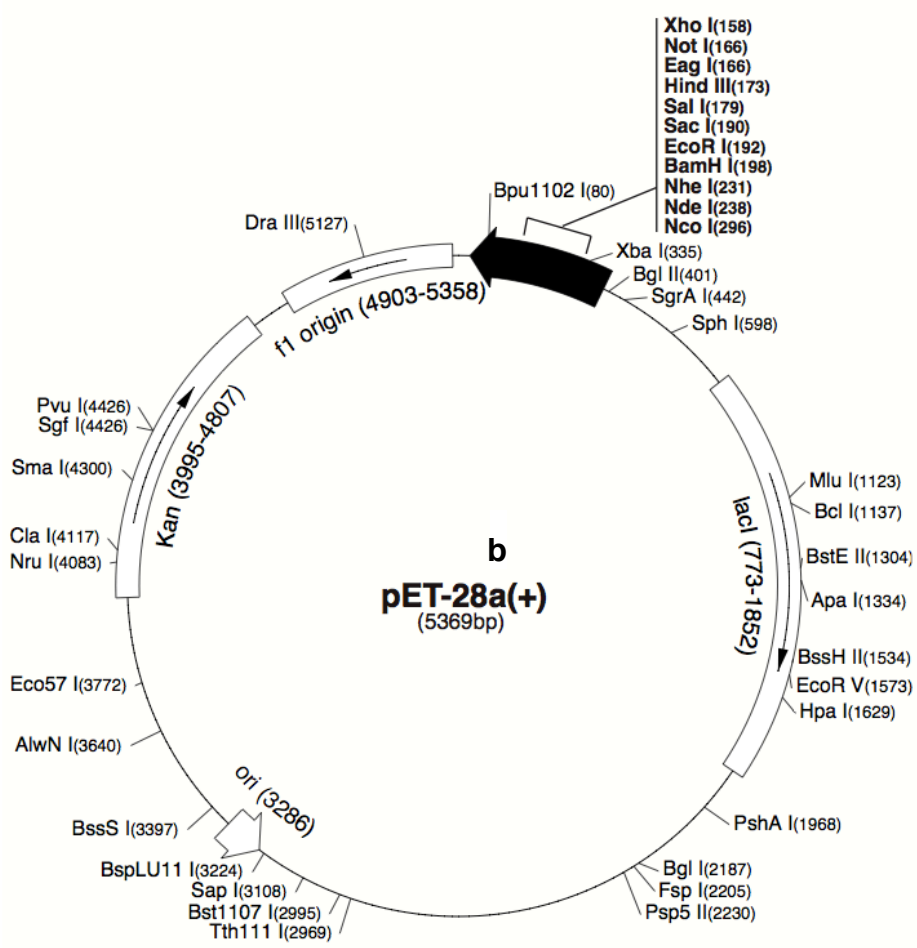


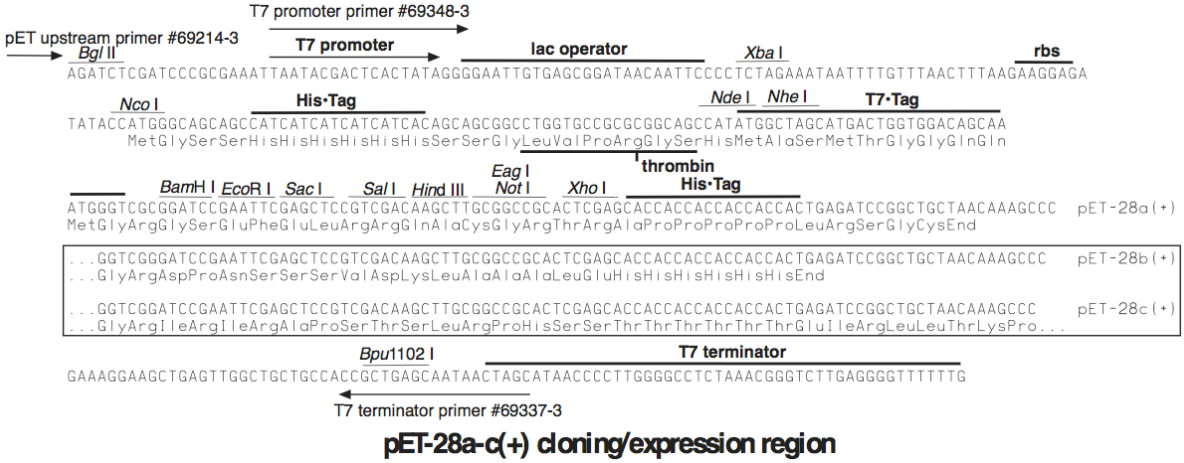


**pET-3a-d cloning/expression region**

APPENDIX 2

## pET-28b




**pET-28a-c(+) cloning/expression region**

APPENDIX 3

$H_6$-proCAT-3

# ProtParam

## User-provided sequence:

```
        10          20          30          40          50          60
MGSSHHHHHH SSGLVPAGSH MFRTFPGIPK WRKTHLTYRI VNYTPDLPKD AVDSAVEKAL

        70          80          90         100         110         120
KVWEEVTPLT FSRLYEGEAD IMISFAVREH GDFYPFDGPG NVLAHAYAPG PGINGDAHFD

       130         140         150         160         170         180
DDEQWTKDTT GTNLFLVAAH EIGHSLGLFH SANTEALMYP LYHSLTDLTR FRLSQDDING

       190
IQSLYGPPPD SPET
```

**Number of amino acids:** 194

**Molecular weight:** 21605.0

**Theoretical pI:** 5.56

**Amino acid composition:**

| | | |
|---|---|---|
| Ala (A) | 14 | 7.2% |
| Arg (R) | 7 | 3.6% |
| Asn (N) | 6 | 3.1% |
| Asp (D) | 15 | 7.7% |
| Cys (C) | 0 | 0.0% |
| Gln (Q) | 3 | 1.5% |
| Glu (E) | 10 | 5.2% |
| Gly (G) | 16 | 8.2% |
| His (H) | 15 | 7.7% |
| Ile (I) | 8 | 4.1% |
| Leu (L) | 17 | 8.8% |
| Lys (K) | 6 | 3.1% |
| Met (M) | 4 | 2.1% |
| Phe (F) | 10 | 5.2% |
| Pro (P) | 15 | 7.7% |
| Ser (S) | 14 | 7.2% |
| Thr (T) | 14 | 7.2% |
| Trp (W) | 3 | 1.5% |
| Tyr (Y) | 8 | 4.1% |
| Val (V) | 9 | 4.6% |
| Pyl (O) | 0 | 0.0% |
| Sec (U) | 0 | 0.0% |

| | | |
|---|---|---|
| (B) | 0 | 0.0% |
| (Z) | 0 | 0.0% |
| (X) | 0 | 0.0% |

**Total number of negatively charged residues (Asp + Glu):** 25
**Total number of positively charged residues (Arg + Lys):** 13

**Atomic composition:**

| | | |
|---|---|---|
| Carbon | C | 974 |
| Hydrogen | H | 1443 |
| Nitrogen | N | 263 |
| Oxygen | O | 290 |
| Sulfur | S | 4 |

**Formula:** $C_{974}H_{1443}N_{263}O_{290}S_4$
**Total number of atoms:** 2974

**Extinction coefficients:**

Extinction coefficients are in units of $M^{-1}$ $cm^{-1}$, at 280 nm measured in water.

| | |
|---|---|
| Ext. coefficient | 28420 |
| Abs 0.1% (=1 g/l) | 1.315 |

**Estimated half-life:**

The N-terminal of the sequence considered is M (Met).

The estimated half-life is: 30 hours (mammalian reticulocytes, in vitro).
>20 hours (yeast, in vivo).
>10 hours (Escherichia coli, in vivo).

**Instability index:**

The instability index (II) is computed to be 26.99
This classifies the protein as stable.

**Aliphatic index:** 70.93

**Grand average of hydropathicity (GRAVY):** -0.449

APPENDIX 4

CAT-3

# ProtParam

## User-provided sequence:

```
        10         20         30         40         50         60
FRTFPGIPKW RKTHLTYRIV NYTPDLPKDA VDSAVEKALK VWEEVTPLTF SRLYEGEADI

        70         80         90        100        110        120
MISFAVREHG DFYPFDGPGN VLAHAYAPGP GINGDAHFDD DEQWTKDTTG TNLFLVAAHE

       130        140        150        160        170
IGHSLGLFHS ANTEALMYPL YHSLTDLTRF RLSQDDINGI QSLYGPPPDS PET
```

**Number of amino acids:** 173

**Molecular weight:** 19395.5

**Theoretical pI:** 4.89

**Amino acid composition:**

| | | |
|---|---|---|
| Ala (A) | 13 | 7.5% |
| Arg (R) | 7 | 4.0% |
| Asn (N) | 6 | 3.5% |
| Asp (D) | 15 | 8.7% |
| Cys (C) | 0 | 0.0% |
| Gln (Q) | 3 | 1.7% |
| Glu (E) | 10 | 5.8% |
| Gly (G) | 13 | 7.5% |
| His (H) | 8 | 4.6% |
| Ile (I) | 8 | 4.6% |
| Leu (L) | 16 | 9.2% |
| Lys (K) | 6 | 3.5% |
| Met (M) | 2 | 1.2% |
| Phe (F) | 10 | 5.8% |
| Pro (P) | 14 | 8.1% |
| Ser (S) | 9 | 5.2% |
| Thr (T) | 14 | 8.1% |
| Trp (W) | 3 | 1.7% |
| Tyr (Y) | 8 | 4.6% |
| Val (V) | 8 | 4.6% |
| Pyl (O) | 0 | 0.0% |
| Sec (U) | 0 | 0.0% |

| | | |
|---|---|---|
| (B) | 0 | 0.0% |
| (Z) | 0 | 0.0% |
| (X) | 0 | 0.0% |

**Total number of negatively charged residues (Asp + Glu):** 25
**Total number of positively charged residues (Arg + Lys):** 13

**Atomic composition:**

| | | |
|---|---|---|
| Carbon | C | 882 |
| Hydrogen | H | 1310 |
| Nitrogen | N | 228 |
| Oxygen | O | 264 |
| Sulfur | S | 2 |

**Formula:** $C_{882}H_{1310}N_{228}O_{264}S_2$
**Total number of atoms:** 2686

**Extinction coefficients:**

Extinction coefficients are in units of $M^{-1}$ $cm^{-1}$, at 280 nm measured in water.

| | |
|---|---|
| Ext. coefficient | 28420 |
| Abs 0.1% (=1 g/l) | 1.465 |

**Estimated half-life:**

The N-terminal of the sequence considered is F (Phe).

The estimated half-life is: 1.1 hours (mammalian reticulocytes, in vitro).
3 min (yeast, in vivo).
2 min (Escherichia coli, in vivo).

**Instability index:**

The instability index (II) is computed to be 24.60
This classifies the protein as stable.

**Aliphatic index:** 75.03

**Grand average of hydropathicity (GRAVY):** -0.414

APPENDIX 5

HPX-1

# ProtParam

## User-provided sequence:

```
        10          20          30          40          50          60
PIGPQTPKAC DSKLTFDAIT TIRGEVMFFK DRFYMRTNPF YPEVELNFIS VFWPQLPNGL

        70          80          90         100         110         120
EAAYEFADRD EVRFFKGNKY WAVQGQNVLH GYPKDIYSSF GFPRTVKHID AALSEENTGK

       130         140         150         160         170         180
TYFFVANKYW RYDEYKRSMD PGYPKMIAHD FPGIGHKVDA VFMKDGFFYF FHGTRQYKFD

       190         200
PKTKRILTLQ KANSWFNCRK N
```

**Number of amino acids:** 201

**Molecular weight:** 23757.0

**Theoretical pI:** 9.26

**Amino acid composition:**

| | | |
|---|---|---|
| Ala (A) | 12 | 6.0% |
| Arg (R) | 11 | 5.5% |
| Asn (N) | 10 | 5.0% |
| Asp (D) | 13 | 6.5% |
| Cys (C) | 2 | 1.0% |
| Gln (Q) | 6 | 3.0% |
| Glu (E) | 9 | 4.5% |
| Gly (G) | 13 | 6.5% |
| His (H) | 5 | 2.5% |
| Ile (I) | 9 | 4.5% |
| Leu (L) | 8 | 4.0% |
| Lys (K) | 18 | 9.0% |
| Met (M) | 5 | 2.5% |
| Phe (F) | 22 | 10.9% |
| Pro (P) | 13 | 6.5% |
| Ser (S) | 7 | 3.5% |
| Thr (T) | 11 | 5.5% |
| Trp (W) | 4 | 2.0% |
| Tyr (Y) | 13 | 6.5% |
| Val (V) | 10 | 5.0% |
| Pyl (O) | 0 | 0.0% |
| Sec (U) | 0 | 0.0% |

| | | |
|---|---|---|
| (B) | 0 | 0.0% |
| (Z) | 0 | 0.0% |
| (X) | 0 | 0.0% |

**Total number of negatively charged residues (Asp + Glu):** 22

**Total number of positively charged residues (Arg + Lys):** 29

**Atomic composition:**

| | | |
|---|---|---|
| Carbon | C | 1105 |
| Hydrogen | H | 1610 |
| Nitrogen | N | 282 |
| Oxygen | O | 293 |
| Sulfur | S | 7 |

**Formula:** $C_{1105}H_{1610}N_{282}O_{293}S_{7}$
**Total number of atoms:** 3297

**Extinction coefficients:**

Extinction coefficients are in units of $M^{-1}$ $cm^{-1}$, at 280 nm measured in water.

| | |
|---|---|
| Ext. coefficient | 41495 |
| Abs 0.1% (=1 g/l) | 1.747, assuming all pairs of Cys residues form cystines |

| | |
|---|---|
| Ext. coefficient | 41370 |
| Abs 0.1% (=1 g/l) | 1.741, assuming all Cys residues are reduced |

**Estimated half-life:**

The N-terminal of the sequence considered is P (Pro).

The estimated half-life is: >20 hours (mammalian reticulocytes, in vitro).
>20 hours (yeast, in vivo).
? (Escherichia coli, in vivo).

**Instability index:**

The instability index (II) is computed to be 35.41
This classifies the protein as stable.

**Aliphatic index:** 53.38

**Grand average of hydropathicity (GRAVY):** -0.587

# APPENDIX 6

## proMMP-1

```
        10          20          30          40          50          60
MFPATLETQE QDVDLVQKYL EKYYNLKNDG RQVEKRRNSG PVVEKLKQMQ EFFGLKVTGK

        70          80          90         100         110         120
PDAETLKVMK QPRCGVPDVA QFVLTEGNPR WEQTHLTYRI ENYTPDLPRA DVDHAIEKAF

       130         140         150         160         170         180
QLWSNVTPLT FTKVSEGQAD IMISFVRGDH RDNSPFDGPG GNLAHAFQPG PGIGGDAHFD

       190         200         210         220         230         240
EDERWTNNFR EYNLHRVAAH ELGHSLGLSH STDIGALMYP SYTFSGDVQL AQDDIDGIQA

       250         260         270         280         290         300
IYGRSQNPVQ PIGPQTPKAC DSKLTFDAIT TIRGEVMFFK DRFYMRTNPF YPEVELNFIS

       310         320         330         340         350         360
VFWPQLPNGL EAAYEFADRD EVRFFKGNKY WAVQGQNVLH GYPKDIYSSF GFPRTVKHID

       370         380         390         400         410         420
AALSEENTGK TYFFVANKYW RYDEYKRSMD PGYPKMIAHD FPGIGHKVDA VFMKDGFFYF

       430         440         450
FHGTRQYKFD PKTKRILTLQ KANSWFNCRK N
```

**Number of amino acids:** 451

**Molecular weight:** 51975.4

**Theoretical pI:** 6.41

**Amino acid composition:**

| | | |
|---|---|---|
| Ala (A) | 27 | 6.0% |
| Arg (R) | 24 | 5.3% |
| Asn (N) | 22 | 4.9% |
| Asp (D) | 33 | 7.3% |
| Cys (C) | 3 | 0.7% |
| Gln (Q) | 23 | 5.1% |
| Glu (E) | 25 | 5.5% |
| Gly (G) | 34 | 7.5% |
| His (H) | 14 | 3.1% |
| Ile (I) | 18 | 4.0% |
| Leu (L) | 27 | 6.0% |
| Lys (K) | 30 | 6.7% |
| Met (M) | 10 | 2.2% |
| Phe (F) | 34 | 7.5% |
| Pro (P) | 28 | 6.2% |
| Ser (S) | 18 | 4.0% |
| Thr (T) | 25 | 5.5% |
| Trp (W) | 7 | 1.6% |
| Tyr (Y) | 22 | 4.9% |
| Val (V) | 27 | 6.0% |
| Pyl (O) | 0 | 0.0% |
| Sec (U) | 0 | 0.0% |

(B)  0  0.0%
(Z)  0  0.0%
(X)  0  0.0%

**Total number of negatively charged residues (Asp + Glu):** 58
**Total number of positively charged residues (Arg + Lys):** 54

**Atomic composition:**

| | | |
|---|---|---|
| Carbon | C | 2356 |
| Hydrogen | H | 3519 |
| Nitrogen | N | 633 |
| Oxygen | O | 678 |
| Sulfur | S | 13 |

**Formula:** $C_{2356}H_{3519}N_{633}O_{678}S_{13}$
**Total number of atoms:** 7199

**Extinction coefficients:**

Extinction coefficients are in units of $M^{-1}$ $cm^{-1}$, at 280 nm measured in water.

Ext. coefficient  71405
Abs 0.1% (=1 g/l)  1.374, assuming all pairs of Cys residues form cystines

Ext. coefficient  71280
Abs 0.1% (=1 g/l)  1.371, assuming all Cys residues are reduced

**Estimated half-life:**

The N-terminal of the sequence considered is M (Met).

The estimated half-life is: 30 hours (mammalian reticulocytes, in vitro).
>20 hours (yeast, in vivo).
>10 hours (Escherichia coli, in vivo).

**Instability index:**

The instability index (II) is computed to be 34.58
This classifies the protein as stable.

**Aliphatic index:** 62.26

**Grand average of hydropathicity (GRAVY):** -0.639

APPENDIX 7

MMP-1

# ProtParam

## User-provided sequence:

```
          10         20         30         40         50         60
FVLTEGNPRW EQTHLTYRIE NYTPDLPRAD VDHAIEKAFQ LWSNVTPLTF TKVSEGQADI

         70         80         90        100        110        120
MISFVRGDHR DNSPFDGPGG NLAHAFQPGP GIGGDAHFDE DERWTNNFRE YNLHRVAAHE

        130        140        150        160        170        180
LGHSLGLSHS TDIGALMYPS YTFSGDVQLA QDDIDGIQAI YGRSQNPVQP IGPQTPKACD

        190        200        210        220        230        240
SKLTFDAITT IRGEVMFFKD RFYMRTNPFY PEVELNFISV FWPQLPNGLE AAYEFADRDE

        250        260        270        280        290        300
VRFFKGNKYW AVQGQNVLHG YPKDIYSSFG FPRTVKHIDA ALSEENTGKT YFFVANKYWR

        310        320        330        340        350        360
YDEYKRSMDP GYPKMIAHDF PGIGHKVDAV FMKDGFFYFF HGTRQYKFDP KTKRILTLQK

        370
ANSWFNCRKN
```

**Number of amino acids:** 370

**Molecular weight:** 42634.7

**Theoretical pI:** 6.17

**Amino acid composition:**

| | | |
|---|---|---|
| Ala (A) | 24 | 6.5% |
| Arg (R) | 20 | 5.4% |
| Asn (N) | 19 | 5.1% |
| Asp (D) | 28 | 7.6% |
| Cys (C) | 2 | 0.5% |
| Gln (Q) | 15 | 4.1% |
| Glu (E) | 18 | 4.9% |
| Gly (G) | 29 | 7.8% |
| His (H) | 14 | 3.8% |
| Ile (I) | 18 | 4.9% |
| Leu (L) | 20 | 5.4% |
| Lys (K) | 20 | 5.4% |
| Met (M) | 7 | 1.9% |
| Phe (F) | 31 | 8.4% |
| Pro (P) | 23 | 6.2% |
| Ser (S) | 17 | 4.6% |
| Thr (T) | 21 | 5.7% |
| Trp (W) | 7 | 1.9% |
| Tyr (Y) | 19 | 5.1% |
| Val (V) | 18 | 4.9% |
| Pyl (O) | 0 | 0.0% |
| Sec (U) | 0 | 0.0% |

(B) 0 0.0%
(Z) 0 0.0%
(X) 0 0.0%

**Total number of negatively charged residues (Asp + Glu):** 46
**Total number of positively charged residues (Arg + Lys):** 40

**Atomic composition:**

| | | |
|---|---|---|
| Carbon | C | 1943 |
| Hydrogen | H | 2853 |
| Nitrogen | N | 519 |
| Oxygen | O | 554 |
| Sulfur | S | 9 |

**Formula:** $C_{1943}H_{2853}N_{519}O_{554}S_9$
**Total number of atoms:** 5878

**Extinction coefficients:**

Extinction coefficients are in units of $M^{-1}$ $cm^{-1}$, at 280 nm measured in water.

Ext. coefficient 66935
Abs 0.1% (=1 g/l) 1.570, assuming all pairs of Cys residues form cystines

Ext. coefficient 66810
Abs 0.1% (=1 g/l) 1.567, assuming all Cys residues are reduced

**Estimated half-life:**

The N-terminal of the sequence considered is F (Phe).

The estimated half-life is: 1.1 hours (mammalian reticulocytes, in vitro).
3 min (yeast, in vivo).
2 min (Escherichia coli, in vivo).

**Instability index:**

The instability index (II) is computed to be 34.71
This classifies the protein as stable.

**Aliphatic index:** 60.65

**Grand average of hydropathicity (GRAVY):** -0.594

APPENDIX 8

$H_6$-proCAT-1

# ProtParam

## User-provided sequence:

```
        10         20         30         40         50         60
MGSSHHHHHH SSGLVPAGSH MFPATLETQE QDVDLVQKYL EKYYNLKNDG RQVEKRRNSG

        70         80         90        100        110        120
PVVEKLKQMQ EFFGLKVTGK PDAETLKVMK QPRCGVPDVA QFVLTEGNPR WEQTHLTYRI

       130        140        150        160        170        180
ENYTPDLPRA DVDHAIEKAF QLWSNVTPLT FTKVSEGQAD IMISFVRGDH RDNSPFDGPG

       190        200        210        220        230        240
GNLAHAFQPG PGIGGDAHFD EDERWTNNFR EYNLHRVAAH ELGHSLGLSH STDIGALMYP

       250        260
SYTFSGDVQL AQDDIDGIQA IYG
```

**Number of amino acids:** 263

**Molecular weight:** 29504.7

**Theoretical pI:** 5.68

**Amino acid composition:**

| | | |
|---|---|---|
| Ala (A) | 16 | 6.1% |
| Arg (R) | 12 | 4.6% |
| Asn (N) | 11 | 4.2% |
| Asp (D) | 20 | 7.6% |
| Cys (C) | 1 | 0.4% |
| Gln (Q) | 15 | 5.7% |
| Glu (E) | 16 | 6.1% |
| Gly (G) | 24 | 9.1% |
| His (H) | 16 | 6.1% |
| Ile (I) | 9 | 3.4% |
| Leu (L) | 20 | 7.6% |
| Lys (K) | 12 | 4.6% |
| Met (M) | 6 | 2.3% |
| Phe (F) | 12 | 4.6% |
| Pro (P) | 15 | 5.7% |
| Ser (S) | 15 | 5.7% |
| Thr (T) | 14 | 5.3% |
| Trp (W) | 3 | 1.1% |
| Tyr (Y) | 9 | 3.4% |
| Val (V) | 17 | 6.5% |
| Pyl (O) | 0 | 0.0% |
| Sec (U) | 0 | 0.0% |
| (B) | 0 | 0.0% |
| (Z) | 0 | 0.0% |
| (X) | 0 | 0.0% |

**Total number of negatively charged residues (Asp + Glu):** 36
**Total number of positively charged residues (Arg + Lys):** 24

**Atomic composition:**

| | | |
|---|---|---|
| Carbon | C | 1305 |
| Hydrogen | H | 1980 |
| Nitrogen | N | 372 |
| Oxygen | O | 400 |
| Sulfur | S | 7 |

**Formula:** $C_{1305}H_{1980}N_{372}O_{400}S_7$
**Total number of atoms:** 4064

**Extinction coefficients:**
Extinction coefficients are in units of $M^{-1}$ $cm^{-1}$, at 280 nm measured in water.

Ext. coefficient 29910
Abs 0.1% (=1 g/l) 1.014, assuming all pairs of Cys residues form cystines

Ext. coefficient 29910
Abs 0.1% (=1 g/l) 1.014, assuming all Cys residues are reduced

**Estimated half-life:**
The N-terminal of the sequence considered is M (Met).

The estimated half-life is: 30 hours (mammalian reticulocytes, in vitro).
>20 hours (yeast, in vivo).
>10 hours (Escherichia coli, in vivo).

**Instability index:**
The instability index (II) is computed to be 31.86
This classifies the protein as stable.

**Aliphatic index:** 67.83
**Grand average of hydropathicity (GRAVY):** -0.664

Back to the Top

APPENDIX 9

Thrombin-cleaved proCAT-1

# ProtParam

## User-provided sequence:

```
        10         20         30         40         50         60
GSHMFPATLE TQEQDVDLVQ KYLEKYYNLK NDGRQVEKRR NSGPVVEKLK QMQEFFGLKV

        70         80         90        100        110        120
TGKPDAETLK VMKQPRCGVP DVAQFVLTEG NPRWEQTHLT YRIENYTPDL PRADVDHAIE

       130        140        150        160        170        180
KAFQLWSNVT PLTFTKVSEG QADIMISFVR GDHRDNSPFD GPGGNLAHAF QPGPGIGGDA

       190        200        210        220        230        240
HFDEDERWTN NFREYNLHRV AAHELGHSLG LSHSTDIGAL MYPSYTFSGD VQLAQDDIDG


IQAIYG
```

**Number of amino acids:** 246

**Molecular weight:** 27707.8

**Theoretical pI:** 5.23

**Amino acid composition:**

| | | |
|---|---|---|
| Ala (A) | 15 | 6.1% |
| Arg (R) | 12 | 4.9% |
| Asn (N) | 11 | 4.5% |
| Asp (D) | 20 | 8.1% |
| Cys (C) | 1 | 0.4% |
| Gln (Q) | 15 | 6.1% |
| Glu (E) | 16 | 6.5% |
| Gly (G) | 22 | 8.9% |
| His (H) | 10 | 4.1% |
| Ile (I) | 9 | 3.7% |
| Leu (L) | 19 | 7.7% |
| Lys (K) | 12 | 4.9% |
| Met (M) | 5 | 2.0% |
| Phe (F) | 12 | 4.9% |
| Pro (P) | 14 | 5.7% |
| Ser (S) | 11 | 4.5% |
| Thr (T) | 14 | 5.7% |
| Trp (W) | 3 | 1.2% |
| Tyr (Y) | 9 | 3.7% |
| Val (V) | 16 | 6.5% |
| Pyl (O) | 0 | 0.0% |
| Sec (U) | 0 | 0.0% |

| | | |
|---|---|---|
| (B) | 0 | 0.0% |
| (Z) | 0 | 0.0% |
| (X) | 0 | 0.0% |

**Total number of negatively charged residues (Asp + Glu):** 36
**Total number of positively charged residues (Arg + Lys):** 24

**Atomic composition:**

| | | |
|---|---|---|
| Carbon | C | 1229 |
| Hydrogen | H | 1871 |
| Nitrogen | N | 343 |
| Oxygen | O | 379 |
| Sulfur | S | 6 |

**Formula:** $C_{1229}H_{1871}N_{343}O_{379}S_6$
**Total number of atoms:** 3828

**Extinction coefficients:**
Extinction coefficients are in units of $M^{-1}$ $cm^{-1}$, at 280 nm measured in water.

Ext. coefficient 29910
Abs 0.1% (=1 g/l) 1.079, assuming all pairs of Cys residues form cystines

Ext. coefficient 29910
Abs 0.1% (=1 g/l) 1.079, assuming all Cys residues are reduced

**Estimated half-life:**
The N-terminal of the sequence considered is G (Gly).
The estimated half-life is: 30 hours (mammalian reticulocytes, in vitro).
>20 hours (yeast, in vivo).
>10 hours (Escherichia coli, in vivo).

**Instability index:**
The instability index (II) is computed to be 30.24
This classifies the protein as stable.

**Aliphatic index:** 69.35
**Grand average of hydropathicity (GRAVY):** -0.657

APPENDIX 10

CAT-1

# ProtParam

## User-provided sequence:

```
         10          20          30          40          50          60
FPATLETQEQ DVDLVQKYLE KYYNLKNDGR QVEKRRNSGP VVEKLKQMQE FFGLKVTGKP

         70          80          90         100         110         120
DAETLKVMKQ PRCGVPDVAQ FVLTEGNPRW EQTHLTYRIE NYTPDLPRAD VDHAIEKAFQ

        130         140         150         160         170         180
LWSNVTPLTF TKVSEGQADI MISFVRGDHR DNSPFDGPGG NLAHAFQPGP GIGGDAHFDE

        190         200         210         220         230         240
DERWTNNFRE YNLHRVAAHE LGHSLGLSHS TDIGALMYPS YTFSGDVQLA QDDIDGIQAI


YG
```

**Number of amino acids:** 242

**Molecular weight:** 27295.3

**Theoretical pI:** 5.15

**Amino acid composition:**

| | | |
|---|---|---|
| Ala (A) | 15 | 6.2% |
| Arg (R) | 12 | 5.0% |
| Asn (N) | 11 | 4.5% |
| Asp (D) | 20 | 8.3% |
| Cys (C) | 1 | 0.4% |
| Gln (Q) | 15 | 6.2% |
| Glu (E) | 16 | 6.6% |
| Gly (G) | 21 | 8.7% |
| His (H) | 9 | 3.7% |
| Ile (I) | 9 | 3.7% |
| Leu (L) | 19 | 7.9% |
| Lys (K) | 12 | 5.0% |
| Met (M) | 4 | 1.7% |
| Phe (F) | 12 | 5.0% |
| Pro (P) | 14 | 5.8% |
| Ser (S) | 10 | 4.1% |
| Thr (T) | 14 | 5.8% |
| Trp (W) | 3 | 1.2% |
| Tyr (Y) | 9 | 3.7% |
| Val (V) | 16 | 6.6% |
| Pyl (O) | 0 | 0.0% |
| Sec (U) | 0 | 0.0% |
| (B) | 0 | 0.0% |
| (Z) | 0 | 0.0% |
| (X) | 0 | 0.0% |

**Total number of negatively charged residues (Asp + Glu):** 36
**Total number of positively charged residues (Arg + Lys):** 24

**Atomic composition:**

| | | |
|---|---|---|
| Carbon | C | 1213 |
| Hydrogen | H | 1847 |
| Nitrogen | N | 337 |
| Oxygen | O | 374 |
| Sulfur | S | 5 |

**Formula:** $C_{1213}H_{1847}N_{337}O_{374}S_5$
**Total number of atoms:** 3776

**Extinction coefficients:**

Extinction coefficients are in units of $M^{-1}$ $cm^{-1}$, at 280 nm measured in water.

Ext. coefficient 29910
Abs 0.1% (=1 g/l) 1.096, assuming all pairs of Cys residues form cystines

Ext. coefficient 29910
Abs 0.1% (=1 g/l) 1.096, assuming all Cys residues are reduced

**Estimated half-life:**

The N-terminal of the sequence considered is F (Phe).

The estimated half-life is: 1.1 hours (mammalian reticulocytes, in vitro).
3 min (yeast, in vivo).
2 min (Escherichia coli, in vivo).

**Instability index:**

The instability index (II) is computed to be 30.58
This classifies the protein as stable.

**Aliphatic index:** 70.50

**Grand average of hydropathicity (GRAVY):** -0.657

Back to the Top

# APPENDIX 11
# PRIMER SEQUENCES

**A219E revertant**
forward: 5'- GTTGCGGCTCAT**GAA**CTCGGCCATTCT -3'
reverse: 5'- AGAATGGCCGAG**TTC**ATGAGCCGCAAC -3'

**R300A**
forward: 5'- GAAGTGATGTTCTTTAAAGAC**GCG**TTCTACATGCGCACAAATCC -3'
reverse: 5'- GGATTTGTGCGCATGTAGAA**CGC**GTCTTTAAAGAACATCACTTC -3'

**F301A**
forward: 5'- GATGTTCTTTAAAGACAGA**GCG**TACATGCGCACAAATCCC -3'
reverse: 5'- GGGATTTGTGCGCATGTA**CGC**TCTGTCTTTAAAGAACATC -3'

**P307A**
forward: 5'- CTACATGCGCACAAAT**GCC**TTCTACCCGGAAG -3'
reverse: 5'- CTTCCGGGTAGAA**GGC**ATTTGTGCGCATGTAG -3'

**F308A**
forward: 5'- CATGCGCACAAATCCC**GCC**TACCCGGAAGTTGAG -3'
reverse: 5'- CTCAACTTCCGGGTA**GGC**GGGATTTGTGCGCATG -3'

**Y309A**
forward: 5'- GCACAAATCCCTTC**GCC**CCGGAAGTTGAGCTC -3'
reverse: 5'- GAGCTCAACTTCCGG**GGC**GAAGGGATTTGTGC -3'

**F316A**
forward: 5'- CGGAAGTTGAGCTCAAT**GCG**ATTTCTGTTTTCTGGCC -3'
reverse: 5'- GGCCAGAAAACAGAAAT**CGC**ATTGAGCTCAACTTCCG -3'

**S243C**
forward: 5'- CCTAGCTACACCTTC**TGT**GGTGATGTTCAGCTA -3'
reverse: 5'- TAGCTGAACATCACC**ACA**GAAGGTGTAGCTAGG -3'

**S318C**
forward: 5'- GAGCTCAATTTCATT**TGT**GTTTTCTGGCCACAACTG -3'
reverse: 5'- CAGTTGTGGCCAGAAA**ACA**CAAATGAAATTGAGCTC -3'

# APPENDIX 12
# EXAMPLE SEQUENCE DATA

proMMP-1* S243C, S318C237